Interaction between Input and Model Uncertainty for Distributed Rainfall-Runoff-Models

Von der
Fakultät Architektur, Bauingenieurwesen und Umweltwissenschaften
der Technischen Universität Carolo-Wilhelmina
zu Braunschweig

zur Erlangung des Grades eines
Doktoringenieurs (Dr.-Ing.)

genehmigte

Dissertation

von
Haytham Shbaita
geboren am 04. März 1979
in Riad, Saudi Arabien

Eingereicht am	26. März 2010
Disputation am	28. Mai 2010
Berichterstatter	Prof. Günter Meon
	PD Dr. Michael Rode
	Prof. Uwe Haberlandt

2011

Contents

Abstract

The work at hand deals with the uncertainty in the hydrological modeling. A subject that is known for a long time, but more recently has been given more attention. The aims of this work are to implement more than one uncertainty analysis method and to compare them, to study the interaction between input (different spatial resolutions and/or rainfall uncertainty) and model (parameter) uncertainties for distributed rainfall-runoff models and to determine the effect of the catchment characteristics and size on the interaction.
For modeling the water balance the model WaSiM-ETH with the TOPMODEL approach was chosen. For quantifying the uncertainty in the modeling three different methods are used, the Bayesian based Monte Carlo Markov Chain (MC^2 with/without autocorrelation) Method, the Generalized Likelihood Uncertainty Estimator (GLUE) and at last a new concept (INPUT) was developed to incorporate input uncertainty in the total uncertainty analysis procedure. Three different catchments were chosen, the Weiße Elster catchment till gauge Gera-Langenberg (2600 km^2), the Weida catchment (99.5 km^2) and the Döllnitz catchment (215 km^2), each with different characteristics (ranging from lower mountains to flat land) and different grid resolutions.
Regardless the uncertainty analysis method (GLUE, MC^2 without autocorrelation or INPUT) used, similar results (Nash Sutcliffe and water balance) were obtained for all three catchments. Using the MC^2 with autocorrelation improved the overall results, and precisely the water balance in the catchment. The reason for this is the fact that the modeling error in previous time step is removed from the next time step. The disadvantage of this method is it can not be used to predict future events. The new concept INPUT allows studying the interaction between rainfall point measurement uncertainty and the parameter uncertainty. The INPUT method promises new insight in the interaction between input and model uncertainties. Though caution should be taken, since the assumptions made about the form of the input uncertainty has a large influence on the uncertainty analysis results. Furthermore no considerable differences between the spatial resolutions for each catchments was observed. For finding the optimal

grid resolution the percentage of information content loss is used as a guideline. A 20% information content loss is tolerated.

Both size and characteristics of the catchment play an important role in the uncertainty analysis. The uncertainty analysis results are as good as the model simulation of the catchment. This is clearly seen for the catchment Döllnitz. This could be considered as a downside of such uncertainty analysis methods, or this raises the difficulty of distinguishing and separating the diverse sources of uncertainty. For the rainfall measurement uncertainty a different behavior is observed. For the small catchment Weida, the rainfall uncertainty plays an important role, on the contrary to the Gera-Langenberg (large catchment), where the rainfall uncertainty does not have any effects at all. For the large catchment the random errors in the rainfall measurements compensate each others effect. Thus in the sum marginal overall effect is achieved.

Uncertainty results should prepared and communicated to the stakeholders. Uncertainty analysis does not end in the research, but it should be treated and made understandable for the authorities, stakeholders and the public.

Zusammenfassung

Hydrologische Modelle und Stoffhaushaltsmodelle sind wichtige Instrumente des Flusseinzugsgebietsmanagements, mit dessen Hilfe Managementmaßnahmen vor der Umsetzung auf ihre Effizienz untersucht werden können. Prognoseunsicherheiten ergeben sich einerseits durch die dem Modell zugrunde liegenden Eingangsdaten, welche in der Natur erhoben werden müssen und somit mit Fehlern behaftet sind. Andererseits handelt es sich bei der Anwendung eines Modells um eine vereinfachte Abbildung der Wirklichkeit und es kann stets nur eine begrenzte Anzahl von Prozessen berücksichtigt werden. Hinzu kommt, dass Wasser- und Stoffhaushaltsmodelle und deren physikalisch basierte Prozessbeschreibungen häufig auf weit größeren Skalen angewendet werden als für welche sie ursprünglich entwickelt wurden. Daher sind Modelle, die für Prognosen eingesetzt werden, ebenfalls eine bedeutende Quelle von Unsicherheiten.

Ziele dieser Arbeit sind:

- die Implementierung und der Vergleich mehrerer Unsicherheitsanalyse-Methoden,
- die Untersuchung der Wechselwirkung zwischen Inputunsicherheiten (verschiedene Gridauflösung und/oder Niederschlagsunsicherheit) und Modell (Parameter)-Unsicherheiten bei der Anwendung von Niederschlags-Abfluss-Modellen und
- die Ermittlung des Einflusses von Gebietscharakteristika und Größe auf die Wechselwirkung.

Für die Modellierung des Wasserhaushalts ist das Modell WaSiM-ETH gewählt worden. Es ist ein gridbasiertes Modell, dass die Klimadaten und den Niederschlag für jede Zelle interpoliert. Die Unsicherheitsanalyse wird mit drei verschiedenen Methoden, Monte Carlo Markov Chain (MC^2, mit/ohne Autokorrelation), Generalized Likelihood Uncertainty Estimator (GLUE) und ein neues Konzept (INPUT), das für diese Arbeit entwickelt wurde, durchgeführt.

Drei unterschiedliche Untersuchungsgebiete wurden gewählt, die Weiße Elster bis Pegel Gera-Langenberg (2600 km^2), die Weida (Mittelgebirge, 99.5 km^2) und die Döllnitz (Loess, 215 km^2). Für jedes Einzugsgebiet wurden die benötigten Eingangsdaten für das Modell WaSiM-ETH (Digitales Höhenmodell, Bodenkarte und Landnutzungskarte) aggregiert, um verschiedene Gridauflösungen (von 100 m bis 6000 m) zu generieren. Für jedes Einzugsgebiet und jede Gridauflösung wurden die drei Unsicherheitsanalyse-Methoden angewendet.

Unabhängig von der Wahl der Unsicherheitsanaylse-Methode (GLUE, MC^2 ohne Autokorrelation oder INPUT) sind die Ergebnisse (Nash Sutcliffe und Wasserhaushalt) für alle drei Einzugsgebiete vergleichbar. Bei der MC^2 mit Autokorrelation wurden die Ergebnisse, besonders der Wasserhaushalt, deutlich verbessert. Der Grund hierfür ist, dass durch die Autokorrelation der Informationsgehalt einzelner Datenpunkte nicht überbewertet wird. Der Nachteil dieser Methode ist, dass sie nicht für Zukunftsprognosen verwendet werden kann. Das neue Konzept INPUT ermöglicht die Quantifizierung der Wechselwirkung zwischen Input- und Modellunsicherheiten. Jedoch werden die Ergebnisse der Unsicherheitsanalyse stark von der Annahme der Form des Inputfehlers beeinflusst. Weiterhin haben die verschiedenen Gridauflösungen nur geringe Auswirkung auf die Modellierungsunsicherheiten. Um die optimale Gridauflösung zu finden, gilt der Richtwert von 20% Verlust des Informationsgehaltes gegenüber der Ausgangsauflösung.

Sowohl Gebietscharakteristik als auch Größe spielen eine wichtige Rolle bei der Unsicherheitsanalyse, deren Ergebnisse die Güte der Modellsimulation wiederspiegeln. Dies muss als Nachteil der Methoden betrachtet werden, da sie nicht in der Lage sind, die Unsicherheitsquellen (außer Parameter- und Gesamtunsicherheiten) zu unterscheiden. Die Ergebnisse der Niederschlagsunsicherheitsanalyse zeigen ein anderes Verhalten. Für das kleinste Einzugsgebiet Weida hat die Niederschlagsunsicherheit eine wichtige Rolle. Im Gegensatz dazu hat sie beim Einzugsgebiet Gera-Langenberg nur geringe Bedeutung, da sich die Niederschlagsmessfehler ausgleichen.

Chapter 1

Introduction

1.1 Problem and Motivation

Since the early beginnings of mankind, water has played and still plays a major roll not only in the daily lifes of the normal people, but also in the rise and fall of civilizations. Huge civilizations and cultures have developed and originated near large rivers and lakes, e.g. Egypt, Babylon and China. Water served not only drinking and irrigation purposes, but also was used as a mean of transport route and for military defence arrangement. Besides all these advantages of living near water, people had also to deal with catastrophic floods and long lasting periods of droughts. Since that time, mankind have tried to understand, to study and especially to predict such processes and events.

With the permanently increasing human population, agricultural lands are intensively operated and even more fertilizers are used to increase the return. During precipitation events fertilizers reach the water bodies. Thus enforcing a chemical change in the water body and thereby a change in the ecological status of the water body. Therefore, the European Union has passed the Water Framework Directive (WFD). It aims at a good ecological status for all natural waters by 2015. In order to accomplish this task, specific measures need to be taken to enhance the chemical and ecological quality of water bodies.
For the assessment of proposed measures and their cost-effectiveness, integrated river basin assessment is indispensable. Hydrological and nutrient transport models are important tools in river basin management and planning. The European Union encourages the member states to support their decision regarding which measurement to implement, using modeling. However, model predictions need to be interpreted with due care as they are always tied up with uncertainty which comprises both input data uncertainty and model uncertainty.

If the uncertainty associated with model results is not explicitly stated, decision makers tend to be overly confident in model results [Lemons (1996); Mowrer (2000); Beven (2008)]. Recognition of uncertainty in model results can serve as an additional criterion for selection of alternative management strategies [Reckhow (1994)]. Therefore, an important aim of uncertainty assessment is to raise awareness of the inherent imperfection of all modelling and the specific uncertainties associated with each modelling exercise. Moreover, uncertainty analysis can serve as a guide for the efficient direction of future work in improvement of data sets and models. Data availability and quality are often serious limitations for the use of distributed hydrological and nutrient transport models. Because data collection is a time-consuming and expensive task, frequently there is not enough information to fully characterize the variability of catchment properties and the available information is tied up with considerable uncertainty that come in the sampling and measurement process as well as in data transformation, interpolation and aggregation.

The two major data format are time series and maps. The first form comprehend such variables as temperature, discharge and rainfall, measured on a certain time basis, hourly, daily or monthly. The uncertainty in the time series arises from e.g. the measurement error and the variability of such variables. Data could also be available in the form of maps or girds, digital elevation model, soil maps or land use. These are characterized through their resolution, the finer the resolution of the map is, in which it was generated, the larger the information content is. For example, taking a 100 km^2 catchment and generating a soil map with a resolution of 100*100 m results in a grid with 10000 cells, using a resolution of 1000*1000 m results in a grid with 100 cells. So the information content 10000 cells has been aggregated to only 100 cells, resulting in a information content loss. This loss is another reason for uncertainty in input data.

On the other hand, models are inherently imperfect as they are always simplified representations of reality and can only take a limited number of processes into account. Also, in hydrological, substance transport and water quality modelling, models are often applied at much larger scales than they were originally developed for, making use of effective parameters. Therefore, models used for prediction are also a source of considerable uncertainty.

The combined uncertainties of input data and models themselves lead to uncertainties in model predictions. However, the contribution of each source of uncertainty can vary greatly for different applications. It is being argued that using increasingly complex model structures may reduce model uncertainty, but at the same time may increase input uncertainty as more detailed data is needed. Although the need for uncertainty assessment in hydrological, transport and water quality modelling has been recognized for a long time [Beck (1987); Beven and

Binley (1992)], uncertainty assessment has gained importance and been widely applied only in recent years and is still not routinely applied in river basin modelling.

1.2 Objectives

The work aims to study the effect of the level of aggregation in the input data on the predictive uncertainty. This will be done mainly using two of the well known methods in uncertainty analysis. The investigation will be caried out in relation to the size of catchment as well as the landscape heterogeneity within catchments. Furthermore, the influence of uncertainty in the time series input data will be also studied. For this a new concept was developed and implemented. The specific objectives of this research are:

- to implement more than one uncertainty analysis method and to compare them.
- to study the interaction between input (different spatial resolutions and/or rainfall uncertainty) and model (parameter) uncertainties for distributed rainfall-runoff models.
- to determine the effect of the catchment characteristics and size on the interaction.

Chapter 2

Uncertainty in Hydrological Modelling

2.1 Definition of Uncertainty

Uncertainty is the deficit in belief in our knowledge. It is not a characteristic of the physical nature, but the result of our incomplete and imprecise perception of it [Goovaerts (1999)]. This deficit is caused through our inability to measure and to comprehend the natural processes. Using mathematical models, such processes in the nature are simulated. Such models are a simplified description of natural processes expressed by mathematical equations. Some models are based on experiments done at the lab scale, and applied at the catchment scale [Watson et al. (1998)]. Based on this definition two sources of uncertainty are identified: input uncertainty and model uncertainty.

One point should be made clear, the difference between error and uncertainty. Error is also the deficit in belief in our knowledge, but an identifiable deficit, which could be measured. Uncertainty is a *potential* deficit. Examples for error in modelling are the round-up error, discretization error and programming error. Merz and Thieken (2005) introduced a new definition of uncertainty. Despite having the same sources of uncertainty, they differentiate between ***natural*** and ***epistemic*** uncertainty. According to their definition, the natural uncertainty results from the variability of the stochastic processes, whereas epistemic uncertainty is due to deficits in knowledge about the process under study. The term natural uncertainty contains quantities, that are temporally as well as spatially variable. Examples for such quantities are the rainfall amount in a certain time period (e.g. 5 years) or the soil distribution in a catchment. The epistemic uncertainty denote our knowledge about the system, how the system is measured, understood and described. The model uncertainty is part of this uncertainty. The advantage of such definition is the differentiation between reducible (epistemic) and non-reducible (natural) uncertainties. For example when forecasting floods, the rainfall amount and its spatial distribution play an important role. Due to the rainfall variability, its amount could not be estimated accurately. Thereby, it is possible to quantify this variability using frequency distributions of the amount and the spatial distribution. Furthermore, the parameters of the frequency distributions are uncertain. In this case, the epistemic uncertainty in the frequency distribution parameters could be reduced by aquiring and using more measurements.

The meaning of the term uncertainty depends on the contexts in which it is being used. For risk assessment uncertainty represents a probability distribution, which is unknown. In the contexts of GIS programs, it means the quantifiable inexactness of a point measurement, whereas the inexactness could be defined as a statistical distribution around a mean value [Mowrer (2000)]

2.2 Sources of Uncertainty

Based on the definitions in section 2.1 two sources of uncertainty are identified: uncertainty which is caused by the model input data (afterwards input uncertainty) and uncertainty due to model parameter (afterwards model uncertainty). Both sources of uncertainty contribute to the overall uncertainty. Figure 2.1 shows the interaction between the different uncertainty sources. The uncertain input data is propagated through the model. There it interacts with the model and its uncertainty sources to reproduce an output. This output is checked against uncertain measured data to see if they fit, i.e. if the model was able to simulate nature.

2.2.1 Input Uncertainty - Spatial Resolution

For hydrological models, numerous studies have shown that model performance depends on spatial discretization (grid cell size or definition and size of hydrological response units) [e.g. Wolock and Price (1994); Bruneau et al. (1995); Refsgaard et al. (1999); FitzHugh and Mackay (2000); Becker and Braun (1999)]. Especially the dependence of the performance of TOPMODEL on spatial discretization of the topographic index has been thoroughly investigated [e.g. Quinn et al. (1991); Zhang and Montgomery (1994)].

Even though the calibrated parameter values differ greatly when using different grid cell sizes with TOPMODEL, this does not necessarily impair model performance as there is a compensating effect between calibrated hydraulic conductivity and grid cell size [Franchini et al. (1996); Brasington and Richards (1998)]. However, some studies found a pronounced breakpoint after which model performance

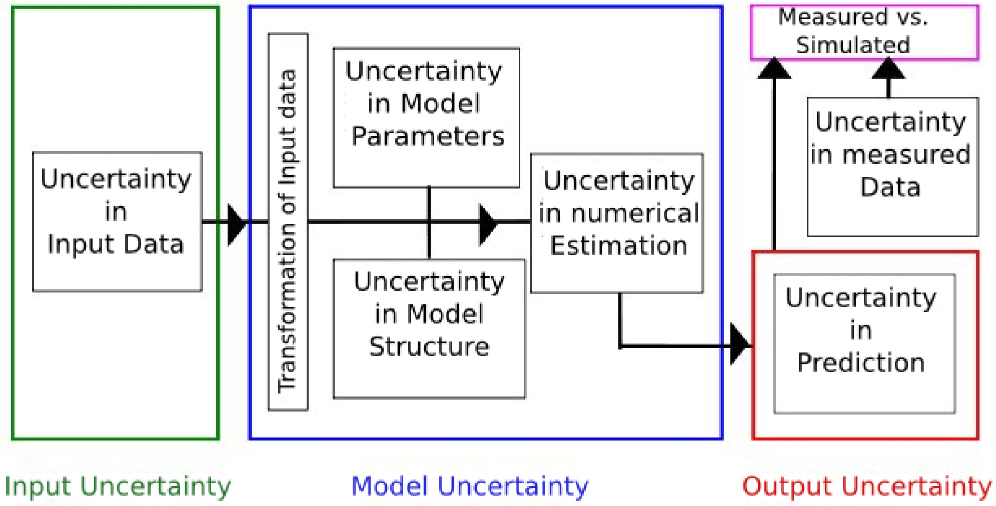


Figure 2.1: Illustration of the different sources of uncertainty and their interaction after Rode (2008).

is severely reduced. This breakpoint was found to be connected to the typical hillslope length in the catchment and could be attributed to the deteriorated information content of the digitial elevation model at grid sizes greater than the average hillslope length [Brasington and Richards (1998)]. Similar results have been presented for other distributed hydrological models [Refsgaard et al. (1999); Vazquez et al. (2002)] though the change in calibrated parameter values was not quite as strong as for TOPMODEL.
Further, it has been shown that the dependence on grid cell size is different for different model outputs, with long term flow volumes being much less sensitive to grid size than the hydrograph shape [Refsgaard et al. (1999)]. Similar studies have been conducted using hydrological responce units (HRU) based distributed modelling [FitzHugh and Mackay (2000)]. They did not find a substantial difference between discharge and sediment yield prediction using 3 to 181 subwatersheds to discreticize a catchment. However, they found substantial differences in sediment generation estimates, showing that there exists a scale dependence, but that in a transport-limited catchment this does not affect sediment yield estimates. Altogether, much less information is available on the scale dependence of nutrient transport models than of hydrological models.
All of these studies are looking at the change of parameter values and model performance (in terms of some performance criterion such as root mean squared error or model efficiency) with a change of spatial scale, but not at the uncertainties associated with the model results, nor at the possible implications for model selection.

2.2.2 Input Uncertainty - Temporal Data

In recent years more and more distributed rainfall-runoff models are being used, with it also the demand for detailed input data. Such data, either spatial or temporal, are in most cases not available. Thus, the used models are usually run with sparse detailed input data. It is assumed that calibrating the model compromises the information loss due to less detailed input data. As explained in the previous section, the input uncertainty results from the spatial variability and our inability to measure natural processes. The most important input data uncertainties sources are

- the climate data and
- the precipitation data.

Since precipitation is the driving force in surface hydrology, it will be handled in a separate section, after a small introduction to the other climate data.

2.2.2.1 Climate Data

Climatology is a branch of meteorology, which handles climate information, the distribution of the climate around the globe, the analysis of the sources and causes of different climates, and the application of climate data to solve special design or operational problems. Climate data comprise precipitation, temperature, global radiation, relative humidity and wind speed and direction [Rode and Wink (2005)].

Temperature is the most measured and documented climate data. The temperature responds sensible and directly to changes in the greenhouse gases. Consequently, the temperature acts as a climate change indicator. Besides soil moisture and global radiation, temperature impacts the vegetation growth, development and reproduction. It could be considered constant over a wide area. For hydrological modelling it is adequate to use daily temperature values. In order to explain 90% of the variation in the daily maximum temperature, a distance of 60 km is required. For the daily minimum temperature a distance of 30 km is needed.

For the uncertainty analysis a normal distribution is assumed for the temperature. The distribution is characterized with a mean being the mean temperature value plus the measurement error, and with a variance equal to the temperature mean value plus the sensor accuracy. For a standard temperature sensor these values are 0.1 ℃ for the measurement error and 1 ℃ for the accuracy [Rode and Wink (2005)].

Global Radiation is an important driver for hydrological processes, leaf and soil temperature. From an energy point of view, most of the global radiation falls in a form of short wave. This short wave global radiation can be measured. In contrast, the long wave are simulated or calculated. The spatial sampling, which gives the mean distance between two installed measurement instruments, varies according to temporal samples needed, between 1 km for actual values, 10 km for daily values and 100 km for monthly mean values. It should be clearly made that the measurements are subjected to the influence of clouds.
Global radiation sensor should be placed at a hight of 3 m. The sky should not be blocked by objects. The sensor has a precision and an accuracy of 5%. For the uncertainty analysis, a normal distribution is taken or a transformation is done before applying the normal distribution [Rode and Wink (2005)].

Relative Humidity or the atmospheric moisture, influences the potential evaporation from the soil and plants. Thus, moisture plays the connecting link between the vegetation and the atmosphere. Due to this, the relative humidity is a driving force for climate change, as well as a response variable. For daily measurements, a distance of 30 km up to 100 km between the climate station is more than enough. Sensors should be able to achieve a precision and an accuracy of ±1%. They should be mounted on a height between 1 m and 2 m. Same as with temperature sensors, they should be mounted in a wooden shelter. For the uncertainty analysis it is important to transform the relative humidity data using an arcsine-square-root function to guaranty the normality [Rode and Wink (2005)].

Wind Speed Besides the soil moisture and the relative humidity, wind speed affects the evaporation rate, and therefor acts as a link between the vegetation and the atmosphere. A distance of 20 km to 100 km between the gauging stations should be sufficient to obtain daily values. The sensor accuracy should have the value of ±10%, and should be mounted on a height of 2 m to 10 m. After many studies, the best distribution to fit the wind speed data is the Weibull distribution [Rode and Wink (2005)]

2.2.2.2 Precipitation

The precipitation includes every sort of water falling down from the sky, rainfall, snow or hail. It is measured by the height of water accumulated over an even area within a certain time or time intervals. Besides that, precipitation plays the main role in surface hydrology, it is also a critical variable in simulating the water balance. For various applications, different demands for the resolution of the precipitation data is needed. For large areas, data with coarse spatial and temporal resolution is required, in contrast to smaller area, where data with high spatial and temporal resolution are necessary. The precipitation has a high temporal and spatial variability. With a distance of 10 km between two rain gauges, only 75% of the precipitation variability can be explained. For 90% a distance of less than 5 km is appropriate. The precipitation station should have the following characteristic:

Precision 0.2 mm

Accuracy ± 1 mm for ≤ 20 mm, $\pm 5\%$ for > 20 mm

Height 1.0 m ± 0.2 m

The main uncertainty sources in regard of the measurement are

1. Gauge Type:
   Gauges differ with respect to construction, form, size and material. Figure 2.2 shows a typical precipitation gauge, which is used by the German Weather Service (DWD). The orifice area varies between 7 cm^2 and 1000 cm^2. Most of the gauges have an area between 100 cm^2 and 200 cm^2. Diverse gauges react differently to weather and other error sources. Furthermore, the loss due to humidification depends on the age, the material of the inside wall and the ratio of the orifice area to the depth of the collector [Rode and Wenk (2006)].

Figure 2.2: A HELLMANN precipitation gauge (Foto: G. Wink).

2. Gauge Height:
   Depending on the country, gauges are mounted at different heights, varying between 0.2 m and 2 m. The higher the gauges are installed, the more they will influence the flow paths of the wind near the orifice. The error induced from the wind depends on the wind speed. Generally, the loss due to wind is about 2% to 15% for rainfall and up to 80% for snow [Rode and Wenk (2006)].

3. Wind Speed and Wind Shield:
   Certain aspects of the gauge design impact the measurement. One of these aspects is the wind shield. It was shown that wind shielded gauges catch up to 50% more precipitation than unshielded. Through a wind shield the wind induced error could be reduced up tp 50% for snow, and up to 70% for mixed precipitation [Rode and Wenk (2006)].

4. Position:
   This affects the gauges that are placed on the hillside. Clearly, gauges placed with the orifice parallel to the hillside are more suitable to catch precipitation as gauges with a horizontal orifice.

5. Inadequate Gauging Grid:
   The error, which due to a faulty design of the gauging grid, is characterized through an insufficient number of gauge and the non-representative location regarding the topography and the geometry of the catchment. Such design errors include:

   a) Biased error of the point measurement of the precipitation.
   b) Biased error of the positioning of the grid.
   c) Distribution of the precipitation through the wind.
   d) Scaling problems.

6. Temperature:
   The temperature affects two phenomena, the moisturizing of the inside wall and the evaporation of water. The first phenomenon describes the retention of water on the walls.

Conducted studies have shown a pattern of highly skewness. In this work, the study from Richter (1995) was used to correct the rainfall data. This will be shown in section 4.3.1.2 page 56.

2.2.3 Model Uncertainty

Models are simplified description of nature. They reflect also our knowledge and comprehension of the processes occurring in nature. Since our knowledge is limited and our methods to measure and catch the natural processes are also limited and imperfect, the models and their prediction are uncertain. This is the first uncertainty source in the model, the simplified description of nature. Other sources are the incomplete identifiability of the model parameters, aggregation of the model results and the over parameterizing of the model. Basically, it is not possible to define all necessary model parameters through measurement [Beven (2001)]. Thus, the model parameters are identified or estimated thought the comparison between the modelled and the measured values, i.e. calibration. Since the distributed rainfall-runoff models have a large number of parameters, different data sets (e.g. discharge, groundwater levels, soil moisture) are needed to calibrate the model. Frequently, such data is not available, and the model parameters are estimated with a minimal data set, and thus, associated with uncertainty.

Most of the uncertainty analysis concentrate on the predictive uncertainty of the model, which is easy to quantify using the parameter uncertainty. The predictive uncertainty deals with the uncertainty in the model prediction. In other words, when using models to predict certain events, the uncertainty in the model prediction is called predictive uncertainty.

The following points summarize the specified sources and types of uncertainties, which could be seen as problematic:

1. Estimating the uncertatiny due to model structure.
2. Incomplete uncertainty analysis due to the model complexity or the uncertainty analysis complexity.
3. Estimating the uncertainty due to biased model error, which grows with time.

In the next sections the two most important sources of model uncertainty, model structure and complexity, will be explained.

2.2.3.1 Model Stucture

As discussed above, the uncertainty analysis deals with model input data and model parameter as the main sources for uncertainties. This also implies the relative easiness, with which such uncertainties could be quantified in comparison to the model structure uncertainty. The model structure uncertainty plays an important role when estimating the model efficiency as well as the parameter uncertainty. Such uncertainties are difficult to estimate and to separate from other uncertainty sources.

The problem of quantifying the model structure uncertainty could be solved with the aid of different process formulations. Refsgaard et al. (2006) have proposed miscellaneous approaches for the quantification of the model structure uncertainty. Their approach could be categorized in two groups according to the availability of measurements for the estimation of the parameters. They show that in most research the case that measurements are available for quantifying the model structure uncertainty is widly spread. This complies with the case of interpolation. Nevertheless, models are applied for extrapolation, in cases where no measurements are available. The authors developed a frame work for quantifying the model structure uncertainty for model applied for extrapolation.

2.2.3.2 Model Complexity

The detail of process description is a key component of model uncertainty. In the last decades, an increasingly large number of ever more complex models have been developed as computing power constantly increased [Beven and Binley (1992)]. As models increase in complexity, they need more detailed data. It is being argued that the data available to date do not support many of the highly complex physically based model structures being developed. This may lead to overparameterisation that is caused by a large number of degrees of freedom (parameters) with not enough detailed input data and independent observations for model calibration and evaluation. According to Van Rompaey and Govers (2002), a more complex model structure will have a lower model error than a simpler one, but due to the increased data demand, input error will be higher. However, most likely this does not hold true for even more complex models as with the increasing number of parameters, their identifiability will decrease if there are not enough observations to support their idenfication (see above), thereby increasing model uncertainty [Janssen and Heuberger (1995)]. Nevertheless, it is important to find a balance between input and model error [Heuvelink (1998), Van Rompaey and Govers (2002)].

In a very extensive comparison study, Perrin et al. (2001) compared 19 lumped rainfall-runoff models with 3 to 9 optimised parameters using data from 429 catchments in different climatic regions in a split sample test. Their results give no indication of more complex models performing better than simpler ones, rather different models perform best in different catchments more or less independent of the number of parameters in the model structure. Also, they find that quite often several models perform practically equally well supporting the concept of equifinality in hydrological modelling.

In a similar study, Gan et al. (1997) argued that the number of (optimised) parameters might not be a good indicator for model complexity as it neglects model structure (nature of mathematic functions, type of interactions). A very comprehensive measure of model complexity has been developed by Snowling and Kramer (2001). Their "index of complexity" is based on the number of parameters and state variables, the number of processes implemented in the model and the sophistication of the mathematical relationships describing these processes and can be calculated using a Petersen matrix, a matrix in which processes comprise the rows and state variables the columns. Refsgaard and Knudsen (1996) compared three models of very different structure: a lumped conceptual (NAM), an intermediate (WATBAL), and a distributed physically-based model (MIKE SHE) in different split sample and proxy basin tests. For gauged catchments, they found no principally better performance by any model. For ungauged catchments the physically-based model performed slightly better.

Van Rompaey and Govers (2002) compared the RUSLE and different simplified versions of it (keeping one parameter constant) in terms of input error which they assessed using Monte Carlo simulation, and model error which they calculated for the simplified versions in relation to the full version of RUSLE. To determine the overall error, they added the MSE of the inputs and the model MSE. They found that when simplifying the model, model error increased, whereas input error decreased. So they determined the "optimal" complexity by comparing the two error sources and found that the simplified version with constant K-factor performed best.

For nutrient fluxes in large scale river basins (Rhine and Elbe), de Wit and Pebesma (2001) compared four different model structures, also looking at model error and input error and at measurement error in the calibration/validation data. They used validation error, measurement error and input error (which they assume to be known) to assess model error and concluded that the available data can only support models of moderate complexity.

VanderPerk (1997) explicitly studied the effect of model structure on accuracy as well as the uncertainty of 8 phosphorous concentration models. He found that with increasing complexity of model structure, model accuracy first increased but then stayed at about the same level, even for more complex models. As the limited calibration data set did not allow for the calibration of some processes, parameter identification decreased with increasing complexity. When correlation between parameters was not taken into account, uncertainty increased as parameter identification decreased, but predictive uncertainty did not change. Thus, he considers a model of medium complexity the most effective as it has the highest accuracy with the least amount of parameters, implying that it has the best parameter identifiability of the more accurate models.

In conclusion it can be noted that there has been considerable research concerning the effect of model complexity (though often described in a rather simplistic way) on the accuracy of the results. Few studies have, however, addressed the change of uncertainty and if so, only very similar model structures were considered. The question of a balance between input and model uncertainty has only rarely been addressed, not at all for distributed hydrological models.

2.3 Uncertainty Analysis Methods

In this sectoin an introduction to various methods in the uncertainty analysis and sensitivity analysis is presented. For the methods used in this study a more detailed overview will be presented in chapter 3.

2.3.1 Monte Carlo Analysis

Monte Carlo (MC) method is a well spread method in uncertainty and sensitivity analysis. This is due to the fact that it is easy to implement and does not require any deep knowledge in mathematics. The term "Monte Carlo" refers to the Monte Carlo Casino in Monaco. Though there are examples for using the Monte Carlo method that goes back to the late nineteenth century (e.g. Buffon's needle), the wide spread use of the method started with the computer era. Nowadays, the Monte Carlo method is applied for the solution and simulation of a wide spectrum of complex problems. In environmental modelling, the earliest implementation of the method dates back to 1968 [Beven (2008)]. In hydrological modelling the method is used to perform a forward uncertainty analysis, sensitivity analysis and to predict parameter uncertainty. There is a similarity between both forward uncertainty analysis and sensitivity analysis. Both sample from a pre-defined distribution in order to explore the model response. One advantage of such method is the capability to define scenarios of boundary conditions. In hydrological modelling this could manifest it self as different land uses scenarios, to study future changes and/or developments. Sensitivity analysis is useful in defining and determining the most sensitive parameters. Thus aiding in reducing the number of parameter for further studies. Monte Carlo method is a way to propagate uncertainty in the input data and/or model parameter through the model and study the effect on the model output. The advantages of such methods are:

1. comprehensive sampling from the pre-defined distribution,
2. model independent ans
3. extensive sampling of model parameters facilitates the identification of non-linearity, thresholds and discontinuity.

A numerous number of algorithms exist for sampling from different types of distributions. The Monte Carlo method is entirely random, meaning, any sample could be taken from the whole range. The term Monte Carlo discribes any method/algorithm which allows a random sampling to obtain an approximate solution of the problem at hand. This is done by

1. choosing the parameters for the study,

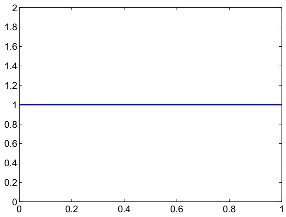
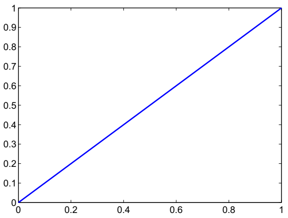
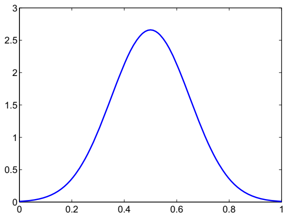
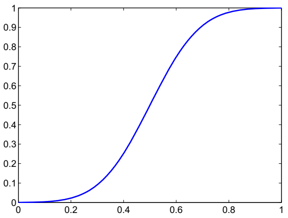
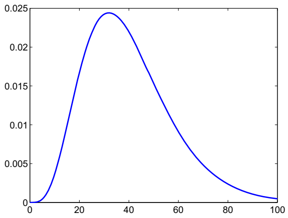
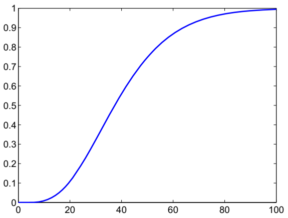
2. defining a certain distribution and
3. randomly sampling from the pre-defined distributions.

This involves deciding on the parameter distributions, sampling from the distribution and then running the model.

2.3.1.1 Parameter Distribution

There are many types of probability distributions, each with its own characteristic and cumulative density function (CDF). For this study, three probability distributions are used. Table 2.1 shows the probability distribution function (PDF) and cumulative density function (CDF). For the uniform distribution a constant probability is assumed over the given range (min and max). This distribution is also called non-informative, since no kind of information flows in the probability function, i.e constant probability for all values within the given range.

Table 2.1: Different probability distribution functions (PDF) and their cumulative density function (CDF)

	Probability Distribution Function (PDF)	Cumulative Density Function (CFD)
Uniform Distribution $p(x) = \alpha$ $x_{min} \leqslant x \leqslant x_{max}$		
Normal Distribution $[\mu = 0.5 \quad \sigma = 0.15]$ $p(x) = \frac{1}{\sqrt{2\pi\sigma_x}} \cdot exp\left[\frac{-1}{\sigma_x}(x-\mu_x)\right]$		
Gamma Distribution $[\alpha = 5 \quad \beta = 8]$ $p(x) = \frac{1}{\alpha\Gamma\beta} \cdot \left(\frac{x}{\alpha}\right)^{\beta-1} \cdot exp\left(\frac{-x}{\alpha}\right)$		

The normal distribution, or Gaussian distribution, on the other hand, assumes that the parameter or the input value is around some mean value (μ) and could change within a certain range or variance (σ). Using this distribution, a strong assumption is made about the parameter or the input value, based on prior knowledge where this value should be. Another point have to be mentioned about this distribution is the fact that it has infinite tails. For values, that should be sampled within a range of [0,1] using a normal distribution with a mean ($\mu = 0.5$) and a variance $\sigma^2 = 0.5$, it is noticed that 95% of the values will be within the range ±2, meaning values smaller than 0 and larger than 1 can be sampled. The gamma distribution does not have this problem, for values of the shape factors α and β larger than 1 it has a unimodal shape.

2.3.1.2 Sampling from the Chosen Distribution

As has been said before Monte Carlo sampling is a random process. Thus, the choice of the pseudo random number generator plays an important role in insuring a certain randomness in the process. To ensure this, the Mersenne Twister pseudo random number generator was used in this work. It was developed in 1997 by Makoto Matsumoto and Takuji Nishimura [Beven (2008)]. It has the advantages:

1. a return period of $2^{19937} - 1$,
2. a very high order of dimensional equidistribution, this implies that there is negligible serial correlation between successive values in the output sequence and
3. it passes almost every test for statistical randomness.

Latin Hypercube Sampling (LHS) Method:
This is a further development in the sampling algorithm. It was programmed so that it uses less number of samples to reproduce the whole distribution. The key point in this algorithm is the breakup of the input ranges into smaller equally wide intervals. Only one sample is taken from each interval to ensure that no two samples are alike. This is the reason why the Latin Hypercube Sampling method is capable of reducing the number of iterations compared to the Monte Carlo method. The number of intervals is the same as the number of iterations. Helton and Davis (2003) showed a 30% reduction in the number of iterations using the Latin Hypercube Sampling over the Monte Carlo method. This method is easy to implement, and guarantees that the samples are taken from the whole distribution.

Importance Sampling Algorithm:
This sampling algorithm belongs to the Monte Carlo Sampling methods. Such methods are developed to overcome the disadvantage of the Monte Carlo method,

which is sampling from the whole parameter interval. According to Kuczera and Parent (1998) the idea behind the importance sampling is to sample from a weighted probability distribution, resembling the posterior probability distribution. This weighted probability distribution is modified iteratively to move forwards the posterior probability distribution. The importance sampling algorithm is characterized through three points:

1. Sampling the parameter values from the uniform distribution and Calculating their probability.
2. Estimating and normalizing the weighting factors

   $$p_i(\theta) = \frac{w_i(\theta)}{\sum\limits_{1}^{N} w_i(\theta)} \tag{2.1}$$

   in this case the weighting factors are the general likelihood function.
3. Upgrading posterior parameter distribution.

In other words, the sampling method divides model space into areas that are of interest (to sample from) and areas that are of less interest, as seen in Figure 2.3.

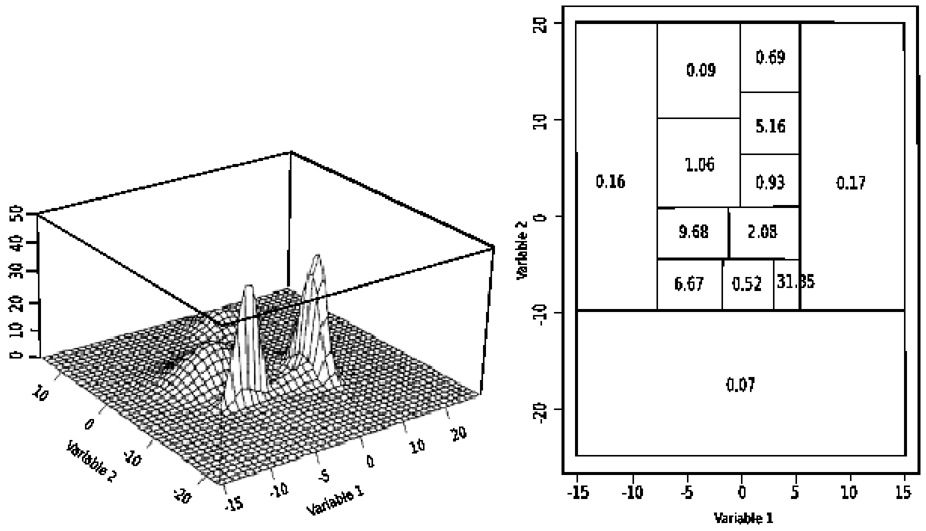


Figure 2.3: Likelihood of a two variable model space, left is the likelihood surface and right are the areas of interest [Spear et al. (1994)].

2.3.2 Differential Analysis

Differential analysis has played an important role in forward uncertainty analysis. This method is based on the partial derivation of the function f with respect to x. The simplest form of differential analysis is the approximation using the Taylor-Series:

$$y(x) = f(x_0) + \sum_{j=1}^{nX} \left[\frac{\partial f(x_0)}{\partial x_j} \right] [x_j - x_{j0}] \tag{2.2}$$

whereas $x_0 = [x_{10}, x_{20}, \ldots, x_{nX,\beta}]$ is a vector with the expected values defined through a certain distribution x_j. After defining the approximation function, the uncertainty in y can be estimated with help of the variance propagation formulas.

$$E(y) = y(x_0) + \sum_{j=1}^{nX} \left[\frac{\partial f(x_0)}{\partial x_j} \right] E[x_j - x_{j0}] \tag{2.3}$$

and

$$V(y) = \sum_{j=1}^{nX} \left[\frac{\partial f(x_0)}{\partial x_j} \right]^2 V(x_j) + 2 \sum_{j=1}^{nX} \sum_{k=j+1}^{nX} \left[\frac{\partial f(x_0)}{\partial x_j} \right] \times \left[\frac{\partial f(x_0)}{\partial x_k} \right] COV(x_j, x_k) \tag{2.4}$$

whereas E is the expected value, V is the variance and COV is the covariance. If the x_j are uncorrelated then

$$V(y) = \sum_{j=1}^{nX} \left[\frac{\partial f(x_0)}{\partial x_j} \right]^2 V(x_j) \tag{2.5}$$

Since the Taylor series is based on the partial derivatives, a sensitivity analysis based on the effect of the individual elements of x on y can be done. Assuming that the elements of x in Equation 2.2 are independent, the fractional contribution of x_j to the variance of y can be approximated by

$$V(y|x_j) = \left[\frac{\partial f(x_0)}{\partial x_j} \right]^2 \frac{V(x_j)}{V(y)} \tag{2.6}$$

with $V(y)$ obtained from Equation 2.5. Ordering the x_j based on the fractional contribution of the total variance, provides an information on how much variance of y can be accounted for by each element of x_j.

The quality of the results is restricted through the Taylor-series approximation, and mainly for non-linear functions. If one would take a first order Taylor approximation, this would lead to a bad representation between the function and its elements [Helton and Davis (2003)]. Since the difficult part of the differential analysis is determining the partial derivatives, most of the research that has been done in recent years was devoted to the development of techniques for the determination of these derivatives.

2.3.3 Response Surface Methodology

The response surface methodology is like the Monte Carlo Analysis, except that the response surface methodology specifies the parameter values not using random sampling but by means of experimental design. Such statistical experimental design can be factorial, fractional factorial, central composite, Plackett-Burman and many more [Helton and Davis (2003)]. Based on the experimental design a set of points for each element in x is chosen. This choice is based only on the experimental design and not on distributions. As a result, there is no probabilistic weight that could be associated with the points, and uncertainty results could not be obtained directly. Instead, an intermediate step is introduced, in which a response surface in the following form

$$y = b_0 + \sum_{j=1}^{nX} b_j x_j \tag{2.7}$$

is constructed. Using the surface within a Monte Carlo simulation the uncertainty in y could be estimated. The response surface in Equation 2.7 is similar to the Taylor approximation in Equation 2.2. So a sensitivity analysis same to the one carried using the Taylor approximation with Equation 2.6 can be performed.

2.3.4 Fourier Amplitude Sensitivity Test and Sobol' Variance Decomposition

The variance of a model $y = f(y)$ could be estimated using the following integration

$$V(y) = \int_{\boldsymbol{\vartheta}_{su}} \left[E(y) - f(x)\right]^2 d_{su}(x) dV_{su} \tag{2.8}$$

whereas $E(y)$ is the expected value

$$E(y) = \int_{\boldsymbol{\vartheta}_{su}} f(y) d_{su} dV_{su} \tag{2.9}$$

Despite the differences in both algorithms, both Fourier Amplitude Sensitivity Test and Sobol' decompose the variance using this formula

$$V(y) = \sum_{1 \leq j \leq nX} V_j + \sum_{1 \leq k \leq nX} V_{jk} + \cdots + V_{1,2,\ldots,nX} \tag{2.10}$$

This is done under the assumption of independence of x_j, where V_j is the variance due to x_j, V_{jk} is the variance due to the interaction of x_j and x_k, and so on. Using the variance decomposition Equation 2.10, different sensitivity measurements could be defined

$$s_j = \frac{V_j}{V(y)} \tag{2.11}$$

$$s_{jk} = \frac{V_{jk}}{V(y)} \tag{2.12}$$

$$s_{jT} = \frac{\left[V_j + \sum_{\substack{1 \leq k \leq l \leq nX \\ \text{k or l=j}}} V_{kl} + \cdots + V_{1,2,\ldots,nX} \right]}{V(y)} \tag{2.13}$$

2.3.4.1 Fourier Amplitude Sensitivity Test (FAST) Method

To estimate the values of $V(y)$ and $E(y)$, the multidimensional integral is converted to a one-dimensional integral using the Fourier transformation. This is done by reducing the n multi-dimensional model (n parameters) to a single dimension, i.e. single parameter. The parameters are transformed into a frequency domain spanned by scaler s as follows

$$c(s) = [G_1(\sin \omega_1 s), G_2(\sin \omega_2 s), \ldots, G_{nX}(\sin \omega_{nX} s)] \tag{2.14}$$

where the G_j are defined functions and the ω_j are defined integers. So the integrals of $E(y)$ and $V(y)$ become

$$E(y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2[c(s)]ds \tag{2.15}$$

and

$$V(y) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f^2[c(s)]ds - E^2(y) \tag{2.16}$$

In general, a numerical integration approach (e.g. Monte Carlo) is usually used to estimate Equations 2.15 and 2.16.
Using Fourier series and its properties the following relations are established:

$$V(y) = \sum_{k=1}^{\infty} \left(A_k^2 + B_k^2 \right) \tag{2.17}$$

where

$$A_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f([c(s)] \cos(ks)de \tag{2.18}$$

$$B_k = \frac{1}{\pi} \int_{-\pi}^{\pi} f([c(s)] \sin(ks)de \tag{2.19}$$

Further more, V_j could be approximated using

$$V_j = \sum_{k=1}^{\infty} \left(A_{k\omega_j}^2 + B_{k\omega_j}^2 \right) \tag{2.20}$$

substituting Equations 2.17 and 2.20 in Equation 2.11 we get

$$s_j = \frac{V_j}{V(y)} = \frac{\sum_{k=1}^{\infty}\left(A_{k\omega_j}^2 + B_{k\omega_j}^2\right)}{\sum_{k=1}^{\infty}\left(A_k^2 + B_k^2\right)} \tag{2.21}$$

2.3.4.2 Sobol' Variance Decomposition

Assuming the model has the form $Y = f(x_1, x_2, .., x_n)$, where $x_1, x_2, \ldots, x_n$ are independent. The joint probability distribution function of the model parameters is

$$P(x_1, x_2, \ldots, x_k) = \prod_{i=1}^{k} p_i(x_i) \tag{2.22}$$

the mean and the variance are

$$E(y) = \int\int \ldots \int f(x_1, x_2, \ldots, x_k) \prod_{i=1}^{k} p_i(x_i) dx_i \tag{2.23}$$

and

$$V(y) = \int\int \ldots \int \{f(x_1, x_2, \ldots, x_k) - E(y)\}^2 \prod_{i=1}^{k} p_i(x_i) dx_i \tag{2.24}$$

$$= \int\int \ldots \int f^2(x_1, x_2, \ldots, x_k) \prod_{i=1}^{k} p_i(x_i) dx_i - E^2(y) \tag{2.25}$$

Fixing one of the parameters to a certain value $x_j = \tilde{x}_j$, results in a variance conditioned on it

$$V(y|x_j = \tilde{x}_j)$$

$$= \int\int \ldots \int \{f(x_1, x_2, \ldots, \tilde{x_j}, \ldots, x_k) - E(y|x_j = \tilde{x}_j)\}^2 \prod_{i=1, i\neq j}^{k} p_i(x_i) dx_i \tag{2.26}$$

$$= \int\int \ldots \int (f^2(x_1, x_2, \ldots, \tilde{x}_j, ..., x_k)) \prod_{i=1, i\neq j}^{k} p_i(x_i) dx_i - E^2(y|x_j = \tilde{x}_j) \tag{2.27}$$

For the purpose of sensitivity analysis one is interested in eliminating the dependence upon the value x_j by integrating $V(y|x_j = \tilde{x_j})$ over the probability density function of $\tilde{x}_j$, by doing this one obtains

$$E(V(y|x_j)) = \int\int ... \int f^2(x_i, x_2, ..., x_j, ...x_k) \prod_{i=1}^{k} p_i(x_i) dx_i - \int E^2(y|x_j = \tilde{x_j}) p_j(\tilde{x}_j) d\tilde{x}_j \tag{2.28}$$

The dependence on $\tilde{x}_j$ is dropped from the left-hand side of Equation 2.28, due to the integration. Subtracting Equation 2.28 from Equation 2.25 one obtains:

$$V(y) - E(V(y|x_j)) = \int E^2(y|x_j = \tilde{x}_j)p_j(\tilde{x}_j)d\tilde{x}_j - E^2(y) \qquad (2.29)$$

Thanks to a statistical identity, the left hand side of Equation 2.29 is also equal to $V(E(y|x_j))$, and represents a measure of the sensitivity of y with

$$S_j = \frac{V(E(y|x_j))}{V(y)} \qquad (2.30)$$

Estimating the integrals is usually done using Monte Carlo sampling methods. The disadvantage of both methods is the high computational cost.

In the coming chapter, a more detailed look will be taken on the chosen uncertainty analysis tools. Furthermore, a new concept for the integration of the rainfall uncertainty in the study will be presented.

Chapter 3

Methods for Uncertainty Analysis

This chapter deals with the uncertainty analysis methodologies implemented in this study. At first, two of the most used methods, the Generalized Likelihood Uncertainty Estimation (GLUE) method and the Monte Carlo Markov Chain method (MC^2), are reviewed and explained in details. In the last part of this chapter, a new concept for integrating the rainfall uncertainty in the overall uncertainty analysis is introduced.

3.1 Generalized Likelihood Uncertainty Estimation Methodology (GLUE)

The Generalized Likelihood Uncertainty Estimation methodology was developed by Beven and Binley (1992). It has a wide application in the environmental modelling, e.g. hydraulic modelling, erosion modelling, ground water modelling and hydrological modelling. The idea arose from the observation that different calibration methods yield different optimal parameter sets. A major point in the methodology is the concept of Equanifility. It states that there is more than one parameter set, that satisfies certain criteria, distributed in the parameter space [Balin (2004)]. This arises from the fact that we can not assume that the model structure is correct, and therefor we need not to find the true model parameters. One advantage of this concept is it takes into account structural and data effects, even if these errors cannot be addressed explicitly. This is done by searching for a set of parameters/models, that simulate the measured data in an acceptable way based on one or more criteria. A choice on which model or parameter set is ‶acceptable″, or in the GLUE terminology ‶behavioral″ needs to be made. A more detailed discussion on this point will follow later. Another advantage of this methodology is that it handles implicitly the modelling errors, without making assumptions on how these errors look like. Such assumptions on the error structure might be wrong or lead to overconditioning. Since no assumption on the form of the error structure is made, model deficiencies are not compensated by an error model. Meaning, when ever the model cannot reproduce the measured data or the system at study, this will manifest itself in prediction limits not being able to bracket the measured data in a satisfiable way.

One disadvantage of this method is that there is no theory to help in making the choice about what kind of likelihood to be used. Another point, which has been mentioned and will be discussed later on is the decision on what is behavioral and what is not. A third disadvantage is without including a formal representation of the error structure, there is no possibility of separating the different sources of error. One point that should be said about formal methods, is that with such methods a separation of the different sources of error might not be very secure, when some sources of error are not included.

The GLUE methodology is conceptually simple, one needs only to make a decision on the following points

- Which informal/formal likelihood measure or measures need to be used to evaluate each model run, and also on the rejection criteria (decision whether the model is behavioral or not)?
- Which model parameter to be included in the uncertainty analysis?
- What kind of prior distribution(s) from which the parameters are to be sampled?
- Which sampling method to be used in generating random realization?

3.1.1 General Likelihood Function

The first step when applying GLUE is the choice of the likelihood function. This method gets its name from the fact that no assumption is taken regarding the structure and the form of the error in the model. Here are some examples of such generalized likelihood functions [Balin (2004)]

Nash Sutcliffe Criterion NS $1 - \frac{\sum\left(Q_{measured} - Q_{simulated}\right)^2}{\sum\left(Q_{measured} - \overline{Q}_{measured}\right)^2}$

Sum of Square Residuals SSR $\sum\left(Q_{measured} - Q_{simulated}\right)^2$

Sum of Square Logarithmic Residuals SSLR

$$\sum\left[\ln(Q_{measured}) - \ln(Q_{simulated})\right]^2$$

Sum of Absolute Error SAE $|Q_{measured} - Q_{simulated}|$

The choice of informal measurements and the implicit handling of errors is still a controversial issue when using GLUE [Montanari (2005), Mantovan and Todini (2006)], this is due to the decision that should be made, whether a model is behavioral or not. It should be mentioned that also formal likelihoods can be used within the GLUE frame work.

3.1.2 Behavioral or not behavioral?

As mentioned before, this is one of the disadvantages of the GLUE methodology, the choice on whether the model is behavioral or not. This decision should be taken before starting the GLUE analysis. It could be the case, that with the chosen criteria no model or parameter set will fulfill the criteria and no model will be rated as behavioral. In this case, the criteria should be relaxed in order to get behavioral models. This could be illustrated by taking a large number of feasible parameter sets from the model space and estimating some likelihood measurement (L), which reflects some belief in each model. Then each parameter value is plotted against the likelihood measurement (L).
Such plots, seen in Figure 3.1, are referred to as dotty plots, where each dot represents a parameter value and the corresponding likelihood function value. Such plots do not give a good impression of the complex interaction between the different parameters, but they clearly show where the parameters associated with the highest likelihoods are. The concept of Equanifility can be clearly seen in Figure 3.1. One notice that regardless of the value taken within the whole range of feasible values, highest values of the likelihood L are achieved. For Parameter 4 values within the range between 10 and 700, obtain the highest likelihood values [Beven (2008); Beven and Binley (1992)].

3.1.3 Meaning of GLUE Prediction Limits

The prediction limits are dependent on the choice of the likelihood. The prediction limits are taken from the quantiles of the prediction over all the behavioral models. The disadvantage of such prediction limits, since no formal error model is used, they will not provide formal estimates of the probability estimating a certain event. A major advantage is the fact that due to the concept of Equanifility, non-stationarities in the residual errors and model failure are more clearly revealed [Beven (2008); Beven and Binley (1992)].

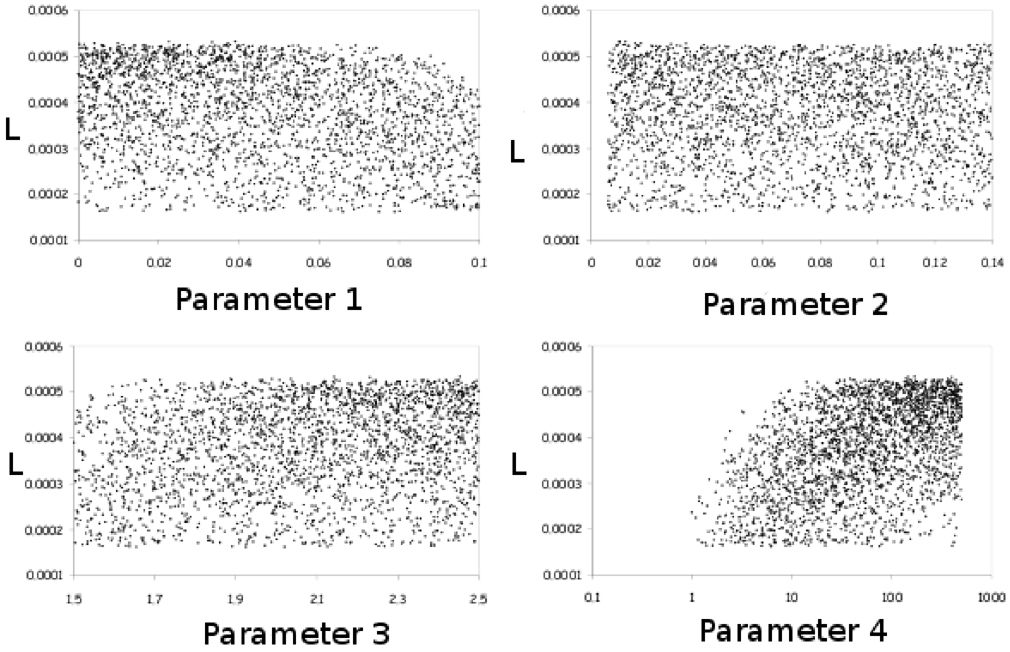


Figure 3.1: Dotty Plots - parameter values against a likelihood measurement (L) after Beven and Binley (1992).

3.2 Bayesian Methods

The origin of the Bayesian methods lies in the paper written by Rev. Thomas Bayes, found among his papers after his death*. Bayesian methods are statistical method, which uses a probability model. Using such methods prior distribution of parameter is modified according to a certain likelihood, leading to a posterior distribution of the parameters. The likelihood function reflects the likelihood or probability of a model to reproduce a certain event. This represents a statistical learning tool. Bayes' theorem could be derived using the conditional probability properties, conditioning the model parameter θ on the data y we get

$$p(\theta|y) = \frac{p(\theta, y)}{p(y)} \tag{3.1}$$

whereas $p(\theta, y)$ is the joint probability of θ and y, $p(\theta|y)$ is the conditioned probability of the parameter θ given y, which represent the posterior distribution. Now conditioning the data y on the parameter θ we get

$$p(y|\theta) = \frac{p(y, \theta)}{p(\theta)} \tag{3.2}$$

---

*A more general discrete form was developed (independently) in France by Pierre-Simon Laplace [Beven (2008)] whereas $p(y,\theta)$ is the joint probability, $p(y,\theta)$ is the likelihood of representing y given the parameter θ. Since the joint probabilities are equal

$$p(\theta,y) = p(y,\theta) \tag{3.3}$$

substituting Equation 3.1 in Equation 3.2 we get

$$p(\theta|y) = \frac{p(y|\theta)\cdot p(\theta)}{p(y)} \tag{3.4}$$

since the probability of an event is constant $p(y) = constant$ we get Bayes theorem

$$p(\theta|y) = \frac{p(y|\theta)\cdot p(\theta)}{C} \tag{3.5}$$

$$p(\theta|y) \sim p(y|\theta)\cdot p(\theta) \tag{3.6}$$

which could be also written as

$$p(\theta|y) \sim p(y|\theta)\cdot p(\theta) \Leftrightarrow posterior \sim likelihood \cdot prior \tag{3.7}$$

So the learning process involves the update of the prior distribution using the likelihood function to achieve the posterior distribution [Balin (2004)].
Therefore, the Bayesian methods are based on three probability concepts:

- the prior probability distribution,
- the formal likelihood function and
- the posterior probability distribution and the Bayes' theorem.

3.2.1 The prior Probability Distribution

Prior probability distributions are basically ***subjective*** probability distributions. The Bayesian method treats the model parameter as a random variable, which could be associated with a subjective probability distribution. The prior distributions are subjective because the available knowledge from the modeller flow in defining the distribution. Such distributions can be informative or non-informative. Vague knowledge about the model parameter should be described with a non-informative distribution, e.g. as uniform distribution.

$$p(\theta) \propto constant \tag{3.8}$$

Another example for non-informative distributions is the case of the unknown variance parameter (σ). For the prior distribution an uniform distribution over the interval $(-\infty, +\infty)$ for $\ln(\sigma)$ is assumed [Balin (2004)].

$$p(\ln(\sigma)) \propto constant \tag{3.9}$$

re-writing $p(\ln(\sigma))$ to $p(\sigma)$ leads to a non-informative distribution for $p(\sigma)$ on the interval $(0, +\infty)$

$$p(\sigma) = p(\ln(\sigma))\left|\frac{\partial \ln(\sigma)}{\partial(\sigma)}\right| = constant \cdot \frac{1}{\sigma} \propto \sigma^{-1} \tag{3.10}$$

3.2.2 Formal Likelihood Function

The likelihood function summarizes the available information in the data. It takes the form $L(\phi)$ and could be estimated when evaluating the probability density function $p(y|\theta)$ with the measured data y. Since the measured data are fixed, the likelihood function is only dependent on the parameter θ. One advantage of such formal likelihood functions is the fact that they are objective. The definition of the likelihood functions is derived from the assumptions made about the sources of uncertainty. Assuming an additive error model

$$Q(x,t) = M(\theta, I, x, t) + \epsilon(x,t) \tag{3.11}$$

where Q is the measured discharge, M is the hydrological model predicted discharge, θ the model parameters and I is the input data [Beven (2008)]. The assumptions which are made here are

- the observed data is well known, and error free and
- the model structure is perfect.

A third assumption, which is made about the error is, it has a normal distribution (with a mean of 0 and an unknown variance) and (for this part) the errors are independent from each other [Balin (2004); Beven (2008)].

$$\epsilon = Q(x,t) - Y(\theta, I, x, t) \sim N(0, \sigma) \tag{3.12}$$

writing the likelihood function for one point

$$L(\epsilon|\theta) = \frac{1}{\sqrt{2 \cdot \sigma}} \cdot exp\left[\frac{1}{2 \cdot \sigma^2} \cdot (\epsilon)^2\right] \tag{3.13}$$

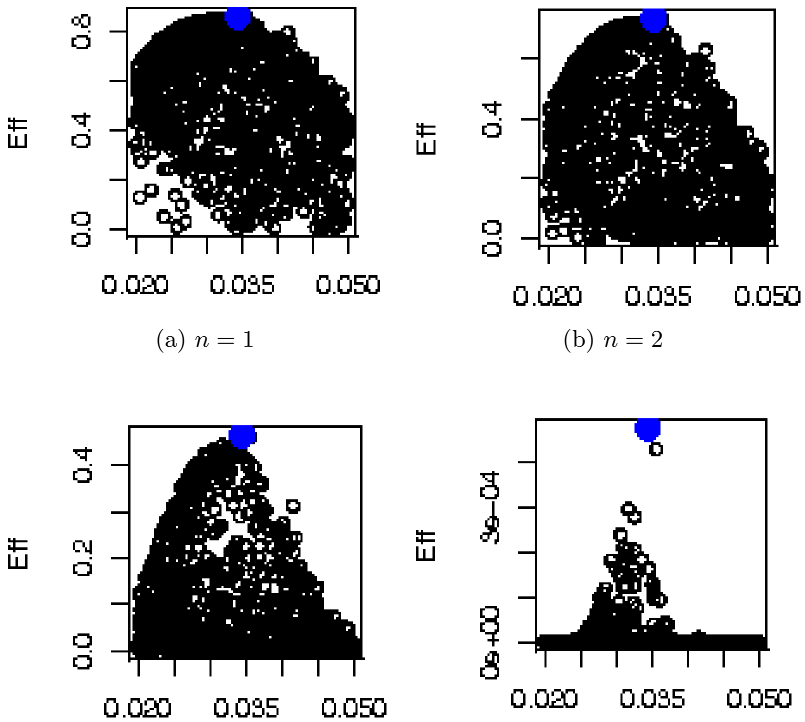
and for n data set

$$L(\epsilon|\theta) = \frac{1}{(\sqrt{2 \cdot \sigma})^n} \cdot exp\left[\frac{1}{2 \cdot \sigma^2} \cdot \sum_{n=1}^{N}(\epsilon)^2\right] \tag{3.14}$$

when dealing with hydrological models, n is in most cases large since it represent the length of the time series. This large n forces the dominant term in Equation 3.14 $[1/(2 \cdot \sigma)^{0.5n}]$ to be large. This means that for models with similar or slight difference in σ will have different likelihoods. This could be clearly seen in Figure 3.2, the larger the value of n, the more spicky is the likelihood function.

Some times the variance of the residuals grows with increasing magnitude of the prediction, this is called *heteroscedasticity*. To over come this problem a Box-Cox-Transformation is introduced. Using the transformation the discharge data is log transformed using the following relation

$$q = \begin{cases} \dfrac{(Q+\alpha)^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln(Q+\alpha) & \lambda = 0 \end{cases} \tag{3.15}$$



(a) $n = 1$

(b) $n = 2$

(c) $n = 5$

(d) $n = 50$

Figure 3.2: Effect of n on the likelihood function.

In this work the Box-Cox parameters were $\lambda = 0$ and $\alpha = 0.0001$. This results in a likelihood function

$$L(\epsilon|\theta) = \frac{1}{(\sqrt{2 \cdot \sigma_{log}})^n} \cdot exp\left[\frac{1}{2 \cdot \sigma_{log}^2} \cdot \sum_{n=1}^{N} (\epsilon_{log})^2\right] \tag{3.16}$$

After dealing with the heteroscedasticity of the residuals, now we concentrate on the correlation in the residuals whether in time or space, which is often found in environmental modelling. Ignoring the correlation in the residual leads to overestimating the information content of each new observation, resulting in over conditioning of the parameter estimates. So if there is a strong correlation between the residuals at different time step, as seen in Figure 3.3, the information content at one time step has already been taken into account in evaluating the model at the previous time step. To eliminate this dependency, a first order autocorrelation is taken into consideration. Using the first order time autocorrelation only the error from the time step $(t-1)$ is removed from the time step (t). This is done using by rewritten equation 3.11 into

$$Q(x,t) = M(I,\theta,x,t) + \epsilon(x,t)_t - AR \cdot \epsilon(x,t)_{t-1} \tag{3.17}$$

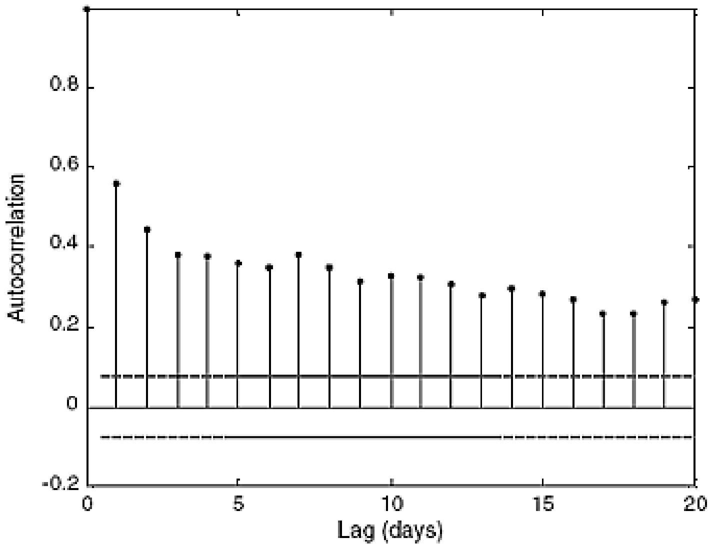


Figure 3.3: Autocorrelation in the residuals of a rainfall-runoff model with a daily time step for the 21000 km^2 Meuse catchment in Belgium, France and the Netherlands using the LISFLOOD model with a 5 km spatial grid [Feyen et al. (2007)].

and the likelihood function becomes

$$L(\epsilon|\theta) = \frac{1}{(\sqrt{2 \cdot \sigma_{log}})^n} \cdot (1 - AR^2)^{0.5} \cdot$$
$$exp\left\{ \frac{1}{2 \cdot \sigma_{log}^2} \cdot \left[(1 - AR^2) \cdot \epsilon_{t,log}^2 + \sum_{n=1}^{N} (\epsilon_{t,log} - AR \cdot \epsilon_{t-1,log})^2 \right] \right\} \tag{3.18}$$

Beside this, the autocorrelation could also act as an indicator of how well the model simulates and capture the physical processes occurring at the catchment scale. A small autocorrelation value indicates that the error at time step (t) has a small dependence on the error from the time step $(t-1)$. Thus, the model is able to simulate catchment in a good way, that the error at step (t) is independent from the error at time step $(t-1)$. On the other hand, a large autocorrelation value (~ 1) reveals a large dependence of the error in time step (t) on the error from the time step before. For an autocorrelation value near 1 implies that almost 100% of error at time step (t) could be explained from the time step before. Thus implying that the model did not capture the physical processes at the catchment scale, despite achieving good results using goodness-of-fit criteria (e.g. Nash Sutcliffe or R^2).

3.2.3 The posterior probability distribution and the Bayes' Theorem

As discussed above, using Bayes' theorem it is possible to update the prior distribution using a likelihood function in a learning process to obtain the posterior distribution. This requires to search the model space to identify the areas of high likelihood, thus the posterior distribution. Such search can be difficult due to the complex high dimensional model space. This has led to the development of search methods and algorithms to characterize the likelihood surface. Monte Carlo Markov Chain methods belong to such methods. In the next section a more detailed view of this method is presented.

3.2.4 Monte Carlo Markov Chain (MC^2) Method

The roots of the Monte Carlo Markov Chain methods dates back to the early fifties in the physical statistics and mechanical statistic. It is used to evaluate the multidimensional integration of the likelihood surface. The most well known Monte Carlo Markov Chain algorithm are the Metropolis-Hasting and a special case of it is the Gibbs within Metropolis-Hasting. The Monte Carlo Markov chains are based on a simple concept. It starts with sampling from a proposal distribution or a transitional kernel for the Markov chain. For each sample the value for the likelihood function is determined by evaluating the model at that point. These points are then used to modify the initial distribution to be closer to the posterior distribution. Another round of model evaluation is done, and the method is checked for convergence to a consistent posterior distribution. If not, another round of samples or iteration is made [Balin (2004)].

The Monte Carlo Markov chain can be divided into two steps

- the choice of the Error function and
- the choice of the sampling method.

The first point involves the choice of the likelihood, which was handled in detail earlier. For the second point first the Metropolis algorithm is explained then the Metropolis-Hasting algorithm, and finally the special case Gibbs within Metropolis-Hasting sampling algorithm [Balin (2004)].

According to Kuczera and Parent (1998) the Metropolis algorithms can be outlined in three points:

- The Jump Rule
- The Acceptance/Rejection Rule
- Monitoring Convergence

The three algorithms differ from each other in the first and the second point, as will be explained later.

3.2.4.1 The Jump Rule

This step is needed to generate a Markov chain, in which the set of samples at step t is dependent on the samples from the previous step $t-1$. There are two types of jumps, or walks, a random walk, where every direction of movement has the same weight or probability, as in Figure 3.4 (a). A forced walk, on the other hand, samples are drawn in the same way as a random walk but in a preferential direction, as in Figure 3.4 (b). For speeding up the convergence of the algorithm, a scaling factor s can be periodically tuned up. The initial value should be set to $2.4/\sqrt{d}$, where d is the number of parameters. The adjustment of this scaling factor should be done in such a way that the acceptance ratio, i.e the number of accepted parameter sets to the total number of parameter set, should vary around 0.23 for $d > 5$ and around 0.43 for $d = 1$ [Balin (2004)].

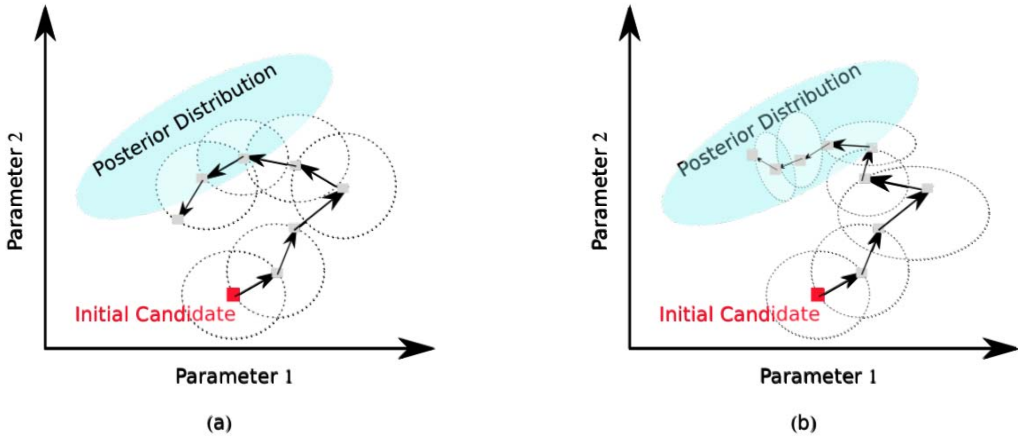


Figure 3.4: Different walk types a random walk (a) and a forced walk (b) within a Markov Chain [Balin (2004)].

3.2.4.2 Acceptance/Rejection Rule

The second step in the Metropolis algorithm is the acceptance or the rejection of the sampled parameter set. First the ratio of the posterior probabilities densities function between the last accepted and the candidate vectors of parameters

$$r = \frac{p(\theta_{new}|Y)}{p(\theta_{old}|Y)} \tag{3.19}$$

whereas θ_{new} is the new candidate and θ_{old} is the last accepted parameter set. The Metropolis Acceptance rule is then applied [Kuczera and Parent (1998), Balin (2004)]

- If $r > 1$ then the new candidate is accepted with a probability of 1

- if $r < 1$ then a random number u in the interval $[0, 1]$ is generated
  - if $r > u$ then the candidate is accepted.
  - If $r < u$ then the candidate is rejected and the last accepted set is taken

The last step is to introduce a random step to check areas with low likelihood in the model space. For the Metropolis-Hasting algorithm only the difference to the Metropolis algorithm is the acceptance rule. It allows asymmetric jump distribution, which could be seen in the acceptance/rejection rule [Balin (2004)]

$$r = \frac{p(\theta_{new}|Y) \cdot p(\theta_{old}|\theta_{new})}{p(\theta_{old}|Y) \cdot p(\theta_{new}|\theta_{old})} \tag{3.20}$$

In this study the third case of the Metropolis algorithm is implemented, Gibbs within Metropolis-Hasting, that is why more attention will be given to this algorithm. Gibbs within Metropolis is a hybrid sampling algorithm. It contains two different special cases of the Metropolis algorithm, the Metropolis-Hasting algorithm and the Gibbs algorithm. The generation of a new parameter set is done using the two algorithms Gibbs and Metropolis. The acceptance rule is defined with the Metropolis-Hasting-Rules. The Gibbs sampler allows to sample a parameter value at once conditioned on all other values, which are kept fixed. In this study the error was assumed to be normal distribution with unknown variance (model variance), the autocorrelation parameter is also introduced in this study. So there exist two extra statistical parameters (σ and AR) besides the model parameter, which need to be estimated.
If the model variance (σ) is only considered, writing the posterior distribution of the parameters given the observed data results

$$(\theta, \sigma^2|Y) \propto L(\theta, \sigma^2) \cdot p((\theta, \sigma^2) \tag{3.21}$$

substituting the likelihood function we get

$$(\theta, \sigma^2|Y) \propto \frac{1}{(\sqrt{2\pi})^n \cdot \sigma^{2 \cdot \frac{n}{2}}} \cdot exp \left\{ -\frac{1}{2} \cdot \frac{\sum_{n=1}^{N} \epsilon^2}{\sigma^2} \right\} \cdot p(\theta, \sigma^2) \tag{3.22}$$

setting

$$S = \frac{\sum_{n=1}^{N} \epsilon^2}{2} \tag{3.23}$$

we get

$$(\theta, \sigma^2|Y) \propto \frac{1}{\sigma^{2 \cdot \frac{n}{2}}} \cdot exp \left\{ -S \cdot \frac{1}{\sigma^2} \right\} \cdot p(\theta, \sigma^2) \tag{3.24}$$

setting $z = 1/\sigma^2$ and we derive the posterior marginal distribution of the model variance σ^2 conditioned on both the observed data y and the model parameters θ we get

$$(\sigma^2|\theta, Y) \propto z^{\frac{n}{2}} \cdot exp\{-S \cdot z\} \tag{3.25}$$

which is a gamma distribution with the parameters $n/2$ and S and is used to sample the model variance after fixing the model parameters. This is called Gibbs within Metropolis-Hasting algorithm. The acceptance/rejection rule is given by Metropolis-Hasting becomes

$$r = \frac{p(\theta_{new}, z_{new}|Y) \cdot p(\theta_{old}, z_{old}|\theta_{new}, z_{new})}{p(\theta_{old}, z_{old}|Y) \cdot p(\theta_{new}, z_{new}|\theta_{old}, z_{old})} \tag{3.26}$$

replacing the posterior distribution by their equivalents, and assuming a uniform distribution for the model parameters, Equation 3.26 becomes

$$r = \frac{L(Y|\theta_{new}, z_{new}) \cdot p(z_{new}) \cdot p(z_{old}|z_{new})}{L(Y|\theta_{old}, z_{old}) \cdot p(z_{old}) \cdot p(z_{new}|z_{old})} \tag{3.27}$$

substituting the corresponding equation and replacing z with $1/\sigma^2$ we get the final form

$$r = exp\left[\left(\frac{1}{\sigma^2_{old}} + \frac{1}{\sigma^2_{new}}\right) \cdot (S_{old} - S_{new})\right] \cdot \left(\frac{S_{new}}{S_{old}}\right)^{\frac{n}{2}} \tag{3.28}$$

3.2.4.3 Monitoring Convergence

A burn-in period is introduced, after which the analysis is done on the chain. The burn-in period is to ensure that the chain has converged and the analysis is done on the part, which represent the posterior distribution. In this study the burn-in period was half of the length of the chain [Balin (2004)].

3.3 Input and Parameter Uncertainty Tool - INPUT

In this section the new developed concept is introduced. The need for this new concept arises from the fact that non of the previous methods handled explicitly the uncertainty in the input data. The background of the new Input and Parameter Uncertainty Tool - INPUT will be explained and then the theory on which it is based.

3.3.1 Background

In this part a new concept will be explained in detail. In this concept the uncertainty in hydrological model is considered differently. Rather than the differentiation between model and input uncertainties, the concept considers epistemic and natural uncertainties. Natural uncertainty is caused by the variability of the system, and epistemic uncertainty results from the ignorance about the system at hand. The natural uncertainty refers to quantities that are inherently spatial and/or temporal variable. Such quantities could be, for example, the precipitation, which is both spatial and temporal variable. Epistemic uncertainties are tied to our capability to measure, describe and understand the system at hand. This split up of the uncertainty sources has the advantage of differentiating between reducible (epistemic) and non-reducible (natural) uncertainties.

A typical uncertainty analysis in the hydrological modelling estimates only either the parameter or the input uncertainty. Recently, different attempts were started to integrate the precipitation into the uncertainty analysis, for example Ajami et al. (2007) and Kavetski et al. (2006). This was done by introducing a correction multiplier for the precipitation, and this multiplier was calibrated. This implies that the precipitation was adjusted till the simulated discharge was accepted. The adjusted or calibrated precipitation does not match the true value. As it has been clearly explained, the disadvantage of this methode is the fact that rainfall is considered as a parameter, which needs to be estimated and calibrated.

Another way to study the interaction of different uncertainty sources, without calibrating the rainfall measurements, is to do a nested Monte Carlo analysis. In the outer loop, the uncertainty due to variability in the precipitation is estimated. In the inner loop, the parameter uncertainty is treaded. The disadvantage of this procedure is the high computational cost. For example if the outer loop has 100 runs and the inner loop 1000 runs, this makes in total $100 \cdot 1000 = 10000$ runs. Normally when doing uncertainty analysis the number of runs starts in the 10000 range. So, if each loop has 10000 runs, this means that the total runs are 10^8 or one hundred million runs. Assuming that each run, takes 1 second to finish, this results in a total run time of 3.17 yeas!.

Hofer et al. (2002) presented a new concept for studying the interaction between the two uncertainty sources. In his concept a simplification for the two nested Monte Carlo analysis is made, in which the two nested loops are transformed to two loops in series. In the next section the concept is explained in detail.

3.3.2 The Theory

Because of the high computational costs of the two-nested Monte Carlo methods (outer: natural and inner: epistemic uncertainty), a new concept will be used in this work to study the effect of natural uncertainties on epistemic uncertainties. Using this method an approximation of the epistemic uncertainty is calculated in the present of the natural uncertainty.

Every model could be presented in this form:

$$Y = h(u^*, V) \tag{3.29}$$

whereas V is the random variable (in this case the precipitation) and u^* is the real but unknown parameter value. Since the parameter values are imprecise known, the resulting state of knowledge will be analyzed. Therefore, the state of knowledge about the u^* will be represented using a subjective probability distribution. Thus Equation 3.29 becomes

$$Y = h(U, V) \tag{3.30}$$

whereas U is a distribution for the unknown parameter values u^*. For each value of Y there will be not only the cumulative distribution $F(Y|u^*)$ but also the value $F(Y|U)$, which has a subjective probability distribution. The concept allows the approximation of the quantiles of the subjective probability distribution for the cumulative distribution $F(Y|u^*)$. Fixing the parameter U to a certain value u, the resulting uncertainties in Y will be caused by the random parameter V. In other words, the conditioned distribution of Y on $U = u$ quantifies the corresponding natural uncertainty in Y. The expected value $E[Y|U = u]$ represents this distribution. Hence, the aim of this concept is to quantify the subjective probability distribution of the conditional expectation $E[Y|U]$. This is done through estimating the expected value and the variance of the conditional expectation, i.e. $E(E[Y|U])$ and $VAR(E[Y|U])$. Applying the probability theory relations

$$E(E[Y|U]) = E(Y) \tag{3.31}$$

$$VAR(E[Y|U]) = VAR(Y) - E(VAR[Y|U]) \tag{3.32}$$

Instead of trying to estimate both $E(E[Y|U]$ and $VAR(E[Y|U])$ values using the complex, computationally costly nested Monte Carlo method, a simplification is introduced to estimate only the three values $E(Y)$, $VAR(Y)$ and $E(VAR[Y|U])$. For this the information from two Monte Carlo runs are needed. In the first run, an approximation for the expected value and the variance of the unconditioned Y are estimated.

$$E(Y) \approx \bar{y} = \frac{1}{n}\sum_i y_i \tag{3.33}$$

$$VAR(Y) \approx \sigma^2 = \frac{1}{n-1}\sum_i (y_i - \bar{y})^2 \tag{3.34}$$

For estimating the third variable $E(VAR[Y|U])$ an assumption is made, the parameter values (U) are fixed to a certain refrence value u_{ref}

$$E(VAR[Y|U]) \approx VAR[Y|U = u_{ref}] \tag{3.35}$$

substituting this assumption in equation 3.32 we get

$$\begin{aligned} VAR(E[Y|U]) &= VAR(Y) - E(VAR[Y|U]) \\ &\approx VAR(Y) - VAR[Y|U = u_{ref}] \end{aligned} \tag{3.36}$$

the reference value of the parameters is estimated in the first Monte Carlo run. The approximation of the distribution $F(Y|U = u_{ref})$ is estimated using the second Monte Carlo run.

$$VAR(Y|U = u_{ref}) \approx \sigma^2 = \frac{1}{n-1} \sum_i (y_i - \bar{y})^2 \tag{3.37}$$

One point should be mentioned here, the results of this concept depend strongly on the assumption made about the rainfall error and its form. Meaning, carefull consideration should be made before assigning a certain distribution to the rainfall error. Another point is, the concept is based on a simplification, which allows to seperate the two nested Monte Carlo runs. This simplification will introduce a certain error / uncertainty in the system, which could be quantified by comparing the results of the two Monte Carlo runs nested and separated.

Chapter 4

Study Catchments and Modelling Tools

4.1 Study Catchments

4.1.1 Weiße Elster

The Weiße Elster catchment with an area of about 5200 km^2 is the largest tributary of the Saale river. From the 718 m elevated spring in Vyhledy (Steingruen) in the Czech Republic, the Weiße Elster river flows through three federal states, Saxony, Saxony-Anhalt and Thuringia until it flows in the Saale river in the south of Halle. It has a length of 250 km, and the major cities of Gera, Leipzig and Halle, and medium sized cities like Zeitz, Plauen and Greiz are located in the river catchment. The major tributaries are the Goeltzsch in the Saxon Vogtland, the Weida in west Thuringia and the Pleiße in the north-west Saxony. The mean discharges of the tributaries are seen in Table 4.1. The river course is highly affected by river straightening measures, specially the lower part of the Weiße Elster river was channalized due to surface mining and land development. In the catchment there are several dams which regulate the water level, provide protec-

Table 4.1: The mean discharges of the River Weiße Elster and its tributaries.

Rever	Area [km^2]	Mean Discharge [m^3/s]
Weiße Elster	5154	25
Parthe	403	1.2
Pleiße	1508	7.9
Weida	459	2.6
Goeltzsch	232	2.8

tion against floods and serve as a source to drinking water. The largest dams are the Pöhl with a volume of 69 Mio. m^3, the Dröda 18 Mio. m^3 and Pirk 10 Mio. m^3.

According to the German Meteorological Service (DWD) the mean temperature changes between 0 ℃ in Fichtelberg and 5 ℃ in Leipzig for the month January. In the summer time (July - August) the temperature ranges between 11 ℃ and 18 ℃. The mean precipitation varies between 832 mm in Bad Elster, 676 mm in Plauen and 570 mm in Leipzig. The land use in the catchment is dominated by agricultural and grassland with 58% of the total area. Other not less important areas are urban areas (16%) and brown coal mining strips (3.1%). Table 4.2 shows the different land uses and their percentage in the catchment. The Weisse Elster catchment covers five different natural landscape units

Table 4.2: Land uses in the Weiße Elster catchment.

Land Use	In Precentage
Water area	1.2 %
Urban Area	15.9 %
Industry	42.6 %
Agriculture	2.6 %
Grassland	15.6 %
Forest	20.6 %
Rest	0.5 %

- Vogtland, west Ore Mountains (Erzgebirge).
- Ore Mountains basin (Erzgebirgsbecken).
- East Thuringia sandstone landscape (Ostthüringer Sandsteinland).
- Loess Hill landscapes (Loesshügelländer).
- Leipziger Landscape, Halle loess hill landscapes and Weißenfels loess hill landscapes.

In Figure 4.1 are the digital elevation model and the different gauging station in the Weiße Elster catchment. Daily data are available for most of the precipitation gauging stations, while six-hourly or hourly data are available for the climate stations. Time series has been collected from 1960 to 2004. A Digital elevation model with a 50 m resolution is available. The land use map was derived from the Landsat imagery at a spatial resolution of 30 m in the year 1999, seen in Figure A.1 page 121. The soil map is based on the Bodenübersichtskarte 1000, with a resolution of 1:1000000, seen in Figure A.2 page 122.

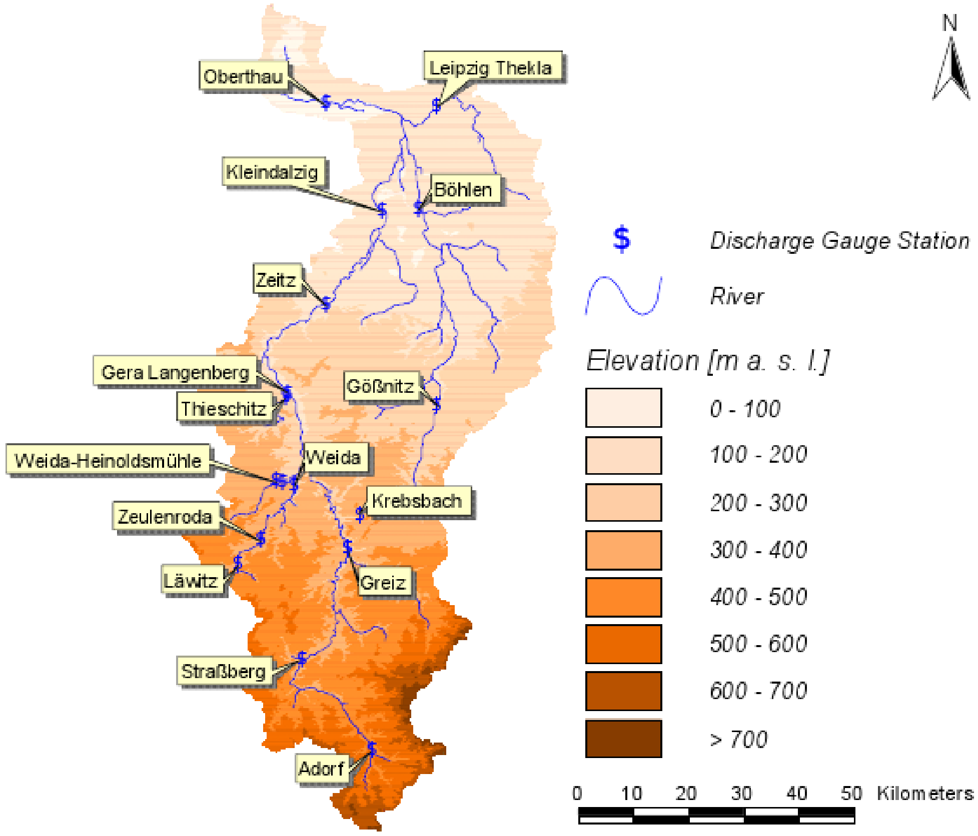


Figure 4.1: The Digital elevation model and the discharge gauges for the Weiße Elster catchment.

In this study not the whole Weiße Elster catchment is taken into consideration, but only up to the gauging station Gera-Langenberg. This reduces the study area to 2600 km^2. The reason for this, is to reduce the computational time needed to do an uncertainty analysis, and at the same time to sustain heterogeneity in the system.

4.1.2 Weida Catchment

The Weida catchment has an area of 99.5 km^2. It is located in the lower mountain range of the federal states Thuringia and Saxony. The Weida catchment, seen in Figure 4.2, is a part of the Weiße Elster catchment. The elevation of the catchment varies between 315 m in the northern part and 565 m in the southern part. Cambisol (Braunerde) and Stagnosol (Pseudogley) are the most common soil types. The land uses are mainly arable land (42%), grass land (25%) and forest (29%).

The climate is moderate and always humid despite the low annual rainfall amount of 690 mm. The majority of rainfall occurs in the summer time. The mean temperature is around 7 ℃. From a hydrological point of view, the catchment is characterized by fast response to precipitation event. The average discharge is 0.72 m^3/s. A highly detailed data base is available. This is with respect to topographical, meteorological data and the discharge data. The time series data have a resolution of one day and go back to 1976. To this time series belong the precipitation from five different stations, the daily mean temperature, the wind speed, the humidity and the global radiation. The topographical data set are the digital elevation model with a resolution of 25 m and the land use map, derived from satellites photos, with a resolution of 30 m, seen in Figure B.1 in page 155. In addition, a soil map exists with 13 different soil types and a resolution of 1:25000, seen in Figure B.2 in page 156.

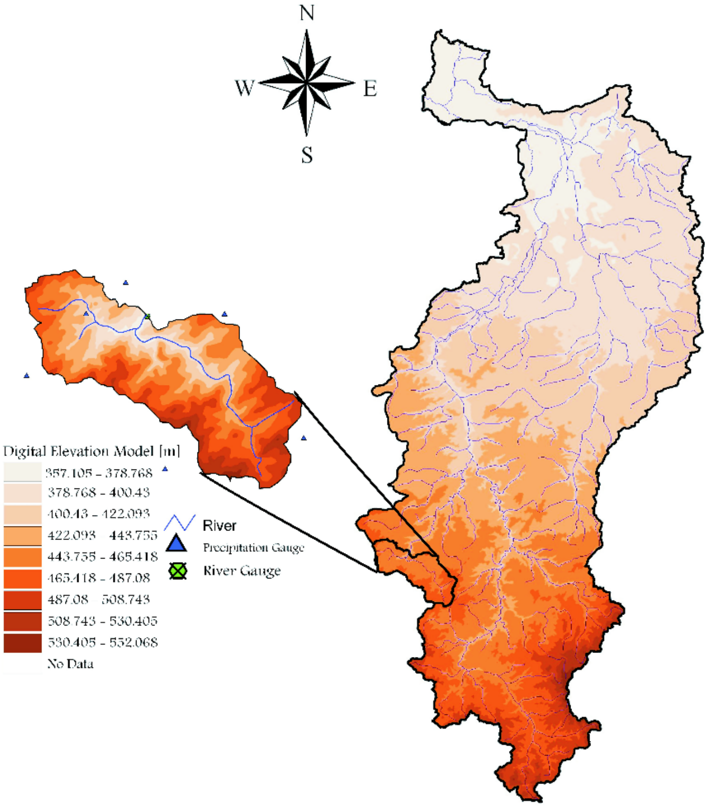


Figure 4.2: Weida catchment as a part of the Weiße Elster catchment.

4.1.3 Döllnitz

The Döllnitz catchment is located in the State of Saxony and flows into the river Elbe at Riesa. It has an area of 215 km^2. In this Study the catchment is considered till the gauge Merzdorf. With a length of 44 km the river Döllnitz flows through the small twons of Mügeln, Thalheim, Oschatz, Borna und Riesa. The main tributaries are the Hasenbach, Grauschwitzbach, Strangraben, Sandbach and the Mühlgraben. The western boarder of the catchment encloses the dams, Horstsee, Göttwitzsee and the Döllnitzsee. Besides the intensive agricultural land use, fishing plays an important role in the western part of the catchment where the dams are located.

The digital elevation model of the catchment could be seen in Figure 4.3. The highest point in the catchment is the Collmberg elevation (312 m above see level) near Collm west from Oschatz near the northern boarder of the catchment. The lowest point is located near the gauge Merzdorf 98 m above see level. The mean elevation is around 170 m. Climatically, the yearly mean precipitation, estimated over the time period from 1990 till 2005 for all precipitation gauges in the catchment, are between 558 $mm/year$ by Riesa (West) and 738 $mm/year$ by Altgeringswalde. The yearly sum of precipitation is 627 $mm/year$. The mean temperature in this period for the station Oschatz is around 9.4 °C.

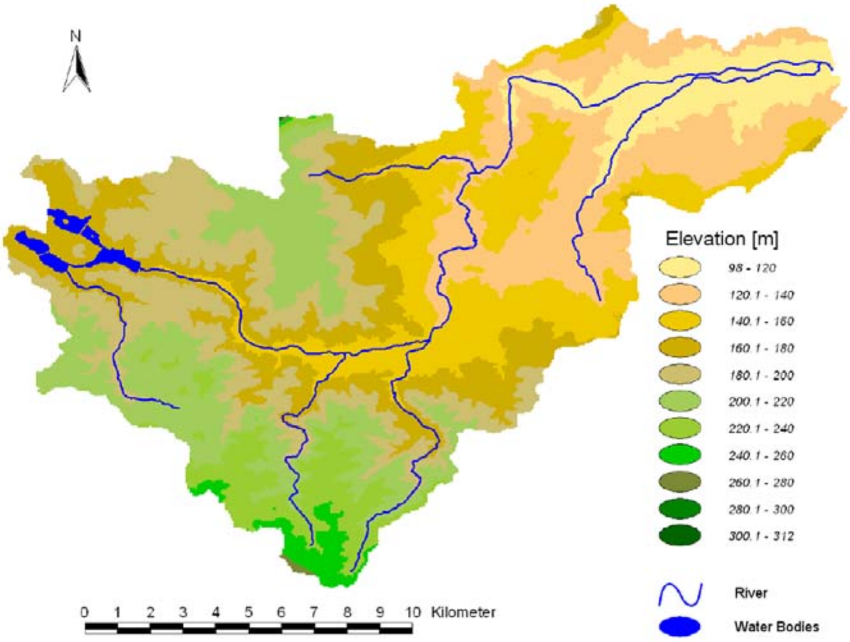


Figure 4.3: The digital elevation model for the Döllnitz catchment.

The climate data for the hydrological modelling of the catchment are provided from the German Meteorological Services (DWD) as daily mean values. For the simulation period, the climate data from the station Oschatz as well as the precipitation data from the four precipitation stations were available. The precipitation data from five stations were available for the modeling. In Table 4.3 is a list of the available stations. The river discharge is measured in two locations seen in Figure 4.4, at the gauging station in Nebitzschen (Area=59.1 km^2) and at the gauging

Table 4.3: Availabe Precipitation and Climate Stations for the Doellnitz Cachment Modeling.

Station Number	Name	Station Art
03377	Oschatz	Climate Station
41324	Obersteina	Precipitation Gauge
41330	Riesa (West)	Precipitation Gauge
41336	Mutzschen	Precipitation Gauge
41265	Altgeringswalde	Precipitation Gauge

station Merzdorf (Area= 211 km^2). The daily discharge values were provided by LfUG (sächsisches Landesamt für Umwelt und Geologie) for the periods from 04.12.2001 till 31.10.2005 (Nebitzschen gauging station) and from 01.11.1911 till 31.10.2005 (Merzdorf gauging station). The gauging station Merzdorf is located 3 km upstream from the confluence point with the River Elbe. Analyzing the discharge data for the gauging station Merzdorf for the period 1912 till 2009, it was also noticed, that an increase in the discharge since 1985 for the period stating from end of August till mid of September. This is due to the controlled drain off of the Döllnitzsee dam located upstream of the main stream. The drain off take place to facilitate fishing. For this a water discharge consent exists, allowing a draw down of a maximal discharge of 1.2 m^3/s over a time period of 21 days before the fishing starts [Hansjürgens and M. (2007)].

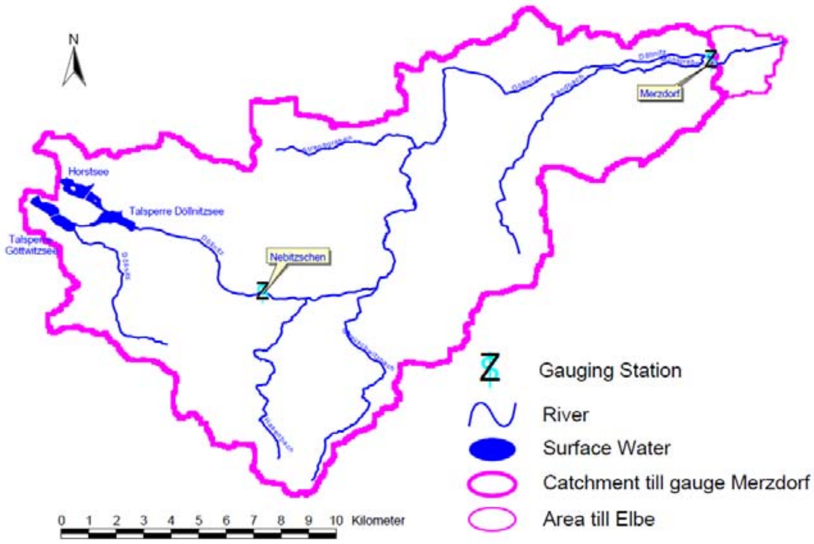


Figure 4.4: The location of discharge gauges in the Döllnitz catchment.

Table 4.4 summarizes the characteristics and the available input data for each catchment.

Table 4.4: Summary of the characteristics of the three Catchments.

	Weida	Döllnitz	Gera-Langenberg
Area	99.5 km^2	215 km^2	2600 km^2
Digital Elevation Model	25 m	50 m	20 m
Land Use	30 m	30 m	30 m
Number of Land Use Categories	13	18	13
Soil Map	1:25000	1:200000	1:1000000
Number of Soil Map Categories	22	12	28
Number of Precipitation Gauges	5	5	49
Number of Climate Stations	1	1	9
Characteristics	Small Scale Lower Mountain	Small Scale Loess Catchment	Meso Scale

4.2 Modelling Tools

4.2.1 Hydrologcial Model - WaSiM

For simulating the hydrological response and the water balance of the catchments, the WaSiM-ETH hydrological model (Water Balance Simulation Model ETH) was used. The model was developed in the late 1990's at the Swiss Federal Institute of Technology in Zurich [Schulla (1997)]. The model is a physically based model, whereas a conceptual approach was implemented when the physical parameters were hard to obtain. For describing the water flow within the unsaturated soil, WaSiM-ETH offers two approaches, a conceptual one TOPMODEL, after Beven and Kirkby (1979), and a physical approach based on the RICHARDS-model for the unsaturated soil coupled to a simple 2D-groundwater flow and transport model for layered aquifers. Both approaches are carried out on a grid, representing the catchment. For each cell in the grid, the meteorological data are also interpolated from the gauging station [Schulla (1997)]. Figure 4.5 shows the structure of the WaSiM-ETH model. WaSiM-ETH is a good established tool for hydrological simulation. It was applied to catchments with areas up to 100000 km^2. As mentioned above, WaSiM-ETH is a grid based model, the following description is only for one cell in the grid, but applies to all cells. The model input data could be divided into two groups:

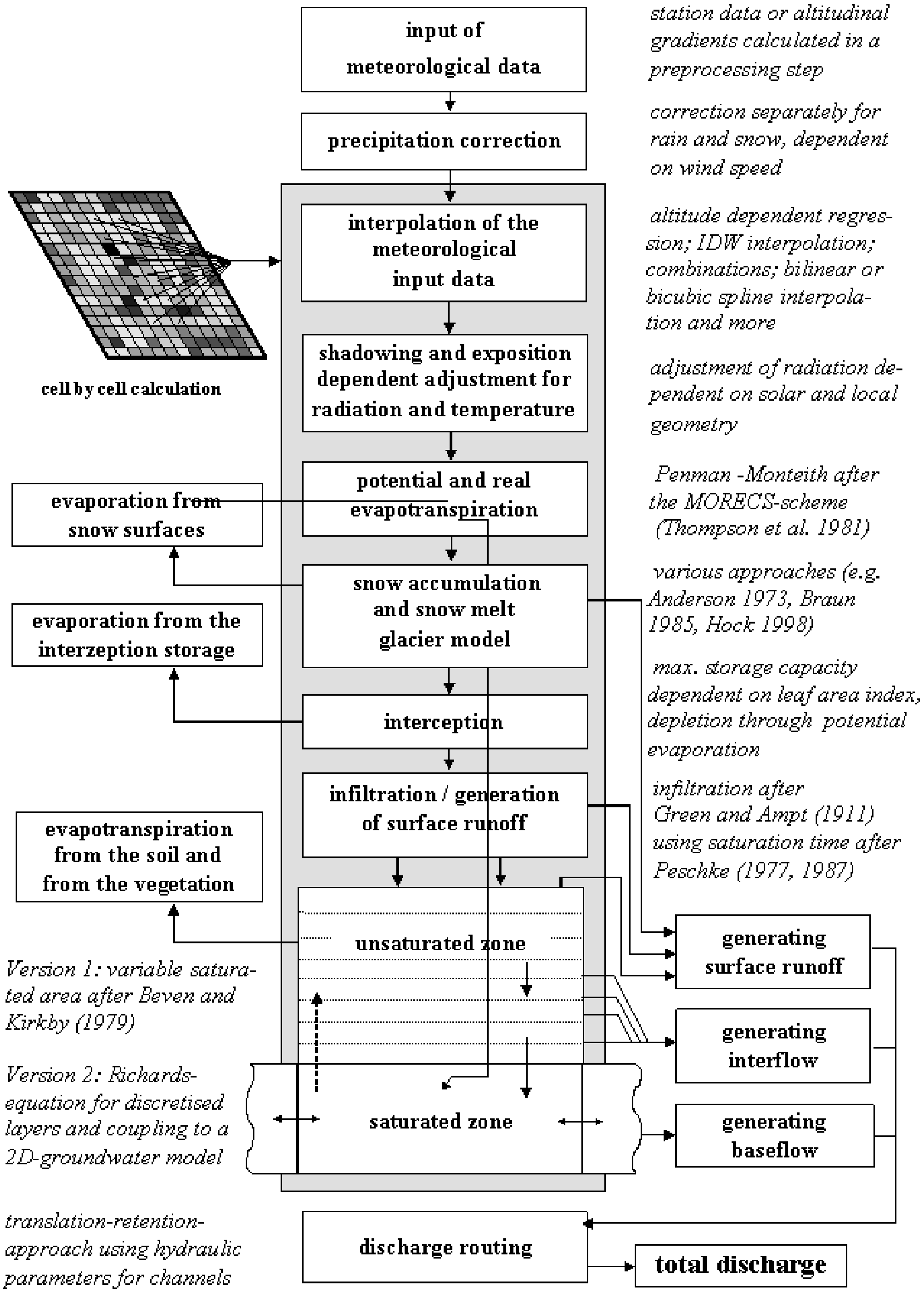


Figure 4.5: Structure of the hydrological model WaSiM-ETH.

- Time Series, and
- Geographical Data

4.2.1.1 The Time Series Group

This contains the precipitation data, the meteorological data, i.e. humidity, global radiation and wind speed, and the discharge data. The digital elevation model, the soil map and the land use map are included in the second group.
The model first corrects precipitation according to the wind speed. This is done for rain and snow separately. Using the inverse distance weighting interpolation method both the precipitation and the meteorological data are interpolated for each cell. In the next step the global radiation and the temperature are corrected depending on the topography and the sun position. With the corrected data the potential evapotranspiration either after Penman-Monteith or Hamon is calculated. The real evapotranspiration is estimated after the TOPMODEL-approach. The snow accumulation and the snow melt are determined after the approach from Anderson (1973); Braun (1985). After Peschke (1977), based on the approach of Green and Ampt (1977), the infiltration is calculated. The discharge from each cell is computed by summing the surplus from the infiltration and the snow melt [Schulla (1997); Schulla J. (1998, 2001)]. In following sections the different modules are explained in more detail.

4.2.1.2 Geographical Data

The geographical data does not only have the function to illustrate the spatial structure of the catchment but also to map the spatial distribution and the constant temporal characteristics. In Table 4.5 are the needed input data for WaSiM-ETH. With the aid of the package *Tanalys*, which is delivered with the model package WaSiM-ETH, it is possible to derive from the digital elevation model the needed input grids for WaSiM-ETH, i.e. topographical index, slope, flow times and flow times sum. Every geographical input must have the same coordinate and spatial dimension.

Table 4.5: The required input data and its derivative for the WaSiM-ETH model.

Data Art	Derived Input for WaSiM-ETH
Digital Elevation Model	Grid, Topographical Index, Slope, Exposition, Flow Times and Flow Times Sum
Land Use	Grid and Parameter Table
Soil Map	Grid, Parameter Table and Transmissivity
Meteorological Data	Interpolated over the whole Catchment

4.2.1.3 Meteorological Data

The model needs, besides the geographical data, the meteorological data, precipitation, temperature, global radiation, wind speed, humidity and the sunshine duration. According to Schulla (1997) daily values of the meteorological data should be sufficient to preform a water balance in a catchment using WaSiM-ETH. A precipitation correction module is implemented in WaSiM-ETH, its function is to adjust uncorrected precipitation data. In this work, the precipitation data were corrected using the correction factors suggested by Richter (1995), hence this module was turned off. The meteorological data are interpolated for each cell in the grid using different interpolation algorithms. The hydrological model offers the following algorithms:

- Altitude Dependent Regression
- Inverse Distance Weighting
  - Inverse Distance Weighting Interpolation
  - Theissen Polygone
- A Combination of the two interpolation methods

In this thesis, the meteorological data were interpolated using the inverse distance weighting interpolation (IDW). The interpolated value for each cell is the sum of all weighted contributing stations data.

$$\hat{z}(u) = \sum_j w_j \cdot z(u_j) \tag{4.1}$$

$$w_j = \frac{1}{d(u,u_j)^p} \cdot \frac{1}{c} \tag{4.2}$$

$$c = \sum_j \frac{1}{d(u,u_j)^p} \text{ which follows} \sum w_j = 1.0 \tag{4.3}$$

whereas

$\hat{z}_u$ the interpolated value at location u

z(u$_j$) observed value at station j

w$_j$ weight of observed value at station j

p weighting power of the inverse distance (between 1 and 3, 2 is recommended)

d(u,u$_j$) distance to the station j

There is the possibility to change the value of the weighting power parameter p and the maximal distance d_{max} as well as two other parameters for fixing the anisotropy [Pöhler (2006); Schulla J. (1998, 2001)].

4.2.1.4 Potential and Real Evapotranspiration

Different approaches are available within WaSiM-ETH to calculate the potential evapotranspiration. These approaches differ in the number of the input parameters, their use and their physical meaning. The work at hand uses the approach from Penman-Monteith to estimate the potential evapotranspiration. This approach requires the temperature, the global radiation and/or the sunshine duration, the relative humidity and the wind speed. The following formula is used to compute the potential evapotranspiration

$$\lambda E = \frac{3.6 \cdot \dfrac{\Delta}{\gamma_P} \cdot (R_n - G) + \dfrac{\rho \cdot c_p}{\gamma \cdot r_a} \cdot (e_s - e) \cdot t_t}{\dfrac{\Delta}{\gamma_P} + 1 + \dfrac{r_s}{r_a}} \tag{4.4}$$

whereas

λ	latent vaporization heat $\lambda = (2500.8 - 2.372 \cdot T)\ kJ/kg$, T: Temperature in °C
E	latent heat flux in $mm/m^2 \equiv kg/m^2$ ($\rightarrow$ [λE]= kJ/m^2)
Δ	tangent of the saturated vapor pressure curve [hPa$\cdot K^{-1}$]
R_N	net radiation
G	soil Heat Flux (here 0.1 R_N) [Wh$\cdot m^{-2}$]
ρ	density of dry air [Kg $\cdot m^{-3}$]
c_p	specific heat capacity of dry air at constant pressure $c_p = 1.005$ [$kJ \cdot (kg \cdot K)^{-1}$]
e_s	saturation vapor pressure at the temperature T [hPa]
e	actual vapor pressure (observed) [hPa]
t_i	number of seconds within a time step
γ_p	psychrometric constant [hPa$\cdot K^{-1}$]
r_s	bulk-surface resistance [s$\cdot m^{-1}$]
r_a	bulk-aerodynamic resistance [s$\cdot m^{-1}$]

Additionally other factors are accounted for, like the specific evaporation potential of the vegetation and the actual soil water content. These factors are simulated using parameters like roughness coefficient, leaf area index, depth of roots and root density distribution as well as different soil parameters.

The real evapotranspiration is estimated depending on the chosen soil module. With the TOPMODEL-approach, which is chosen in this work, the real evapotranspiration is approximated by reducing the potential evapotranspiration. This is done when the soil water content falls below a certain value [Pöhler (2006); Schulla J. (1998, 2001)].

4.2.1.5 Snow Accumulation, Snow Melt and Snow Cover Discharge

In this module the snow accumulation and melt and the resulting discharge is simulated. The snow percentage in the precipitation is determined on the basis of a critical air temperature, at which 50% from the precipitation is snow. The falling snow accumulates to a snow cover. The calculation of the snow melt is done using the temperature-index method [Pöhler (2006); Schulla J. (1998, 2001)].

4.2.1.6 Interception

Interception is the storage of rainfall and melted water is on the vegetation and the ground surface. It is simulated in WaSiM-ETH with a simple bucket approach. The result from the snow module will flow in this module. Therefore the melted water will be directed to the interception storage. The interception capacity depends on the plant type, it's leaf area index, the vegetation coverage degree and the maximum height of water on the leafs. The stored water in the interception storage is removed only through evapotranspiration. If the interception storage is filled, through melt water and/or precipitation, any further water input will infiltrate the soil. This approach assumes an equal thick layer of water over the surface [Pöhler (2006); Schulla J. (1998, 2001)].

4.2.1.7 Infiltration

The infiltration module in WaSiM-ETH is based on the approach from Green and Ampt (1977) with modification from Dyck and Peschke (1995). It requires the initial soil water content and the saturated hydrological conductivity. The approach assumes a homogeneous and unlayered soil structure. In this consideration, the matrix flow plays the major role compared to the micro pore flow. The infiltration process is divided into two phases: the saturation phase and the decrease phase. The infiltrating water is simulated in the module as an advancing wet front. The speed of the front depends on the saturation deficit, the saturated hydrological conductivity and the suction at the front. This module provides the part of the precipitation, which flows as surface runoff, and the other part which infiltrate the soil is considered as an input for the soil model [Pöhler (2006); Schulla J. (1998, 2001)].

4.2.1.8 Soil Model

For the work at hand, the conceptual TOPMODEL-approach was chosen. This a modified version for the soil water balance based on the approach from Beven and Kirkby (1979) extended by capillary rise and interflow. Under the consideration of the topographical index for potential saturated areas the distribution of the saturation deficit can be calculated using the following equations

$$c_s = \ln\left(\frac{\alpha_t}{T_0 \cdot \tan(\beta_t)}\right) \tag{4.5}$$

whereas

c_s topographic index

α_t specific catchment area per unit length of a grid cell; this is the area draining through one meter of the edge of a grid cell $[m^2 \cdot m^{-1}]$

T_0 saturated local hydraulic transmissivity $[m^2 \cdot s^{-1}]$

β_t slope angel $[m \cdot m^{-1}]$

and

$$S_i = S_m - m \cdot \ln\left(\frac{\alpha_t}{T_0 \cdot T_{korr} \cdot \tan(\beta_t)} - \gamma\right) \tag{4.6}$$

with

S_m mean saturation deficit of the basin (arithmetic average of all S_i) $[mm]$

m model parameter $[mm]$

γ mean topographic index of the (sub-)catchment

negative and zero values of S_i produce, in case of rainfall, surface runoff. Should there be a snow layer, part of the snow melt water, which is defined by the user, will flow as surface runoff. Depending on a certain precipitation intensity, the rainfall will be split upon the interception storage and the infiltration. Is the interception storage filled, the excess water will flow as surface runoff [Pöhler (2006); Schulla J. (1998, 2001)].

4.2.1.9 Flow Concentration and Discharge Routing

The sum of the snow melt runoff, the infiltration excess water runoff and the surface runoff build the interflow and the total surface runoff for each cell. The base flow is averaged over the whole (sub-)catchment. The average catchment interflow is the mean of every interflow component in each cell in the grid. Depending on the topographical slope, the direct drain flows from the origin cell to the catchment outlet. This flow concentration is regulated by the slope, the surface roughness, water network concentration and the catchment area. The time lag between the origin of the surface runoff till passing the catchment outlet as well as the flattening of the hydrograph through retention effects are taken into consideration. The flow time for each cell could be calculated with the help of the tool *Tanalys*. Cells with the same flow time are grouped together. The resulting surface runoff in a certain time zone will be forwarded to the time zone underneath it at the beginning of each time step. The total surface runoff is the sum of the catchment interflow, the base flow, the surface runoff in the flow time zone, which after calculating the retention as direct discharge is considered [Pöhler (2006); Schulla J. (1998, 2001)].

4.2.2 Calibration Tool - PEST

A calibration tool aims at matching the model output to the measured variable(s), from nature or the lab. This is done with the aid of an objective function, which describes a relation between the measured and the simulated variable(s). Such relation could be the error. In this case the calibration tool has to minimize the error between the measured and the simulated variable(s). The objective function also plays a role as an indicator of how good the model fits the process at study. The calibration tool changes the parameter of the model, in a certain range given/defined by the user, until the best fit is established. A good calibration tool, after [Skahill and Doherty (2006)], should contain the following points

1. The ability to supply information about the existence and the place of the minima of the objective function.
2. The possibility to provide information about the parameter uncertainty.
3. Should be robust against parameter correlation and ruggedness.
4. Should minimize the number of runs to reach the minimum.

PEST (Parameter Estimation Tool) is a model independent calibration tool. It is Based on the Gauss-Marquardt-Levenberg algorithm. This algorithm has the advantages from both Gauss-Newton method and the Steep Descend method, meaning a less number of runs. The main disadvantage of this method is the fact that the algorithm gets stuck in local minimum, and reports the results to be the global solution. PEST also offers other calibration methods, like the Shuffle Complex Evolutionary (SCE) Method, or a hybrid method between the two mentioned methods, which allows PEST to find the global solution by using a multiple starting points.

The rainfall-runoff model could be represented as a function M, which relates the parameter set p_0 to the runoff q_0.

$$q_0 = M(p_0) \tag{4.7}$$

In the calibration process, the parameters values are changed until the objective function reaches a minimum, or the difference between the measured and modeled discharge is minimum [Doherty (2004)]. PEST has the advantage of a fast convergence, thus finding the optimum in fewer runs compared to other methods. To overcome the disadvantage of the method, getting stuck in local minimum, a new development was implemented. This allows using multiple starting points to find the global solution. Using the Multiple Starting Points method, multiple calibration chains with different starting parameters values are initiated. At the end the best of these chains is taken to be the global minimum. As mentioned before, the Shuffle Complex Evolutionary method from Duan et al. (1992) is another calibration method. The advantage of this method is that it scans the whole parameter space, the disadvantage is, it demands a large number of runs, and thus is highly costly in terms of computational demands [Doherty (2004)].

4.3 Procedure

In this section the procedure to conduct the calibration and the uncertainty analysis is explained. This section is divided into subsections depending on the method at hand. At the beginning the choice of the model parameter, to be considered in both the calibration and the uncertainty analysis study is made. The rainfall correction is then explained. For model calibration, the different objective functions are introduced. Afterwords, the calibration of the model WaSiM-ETH for each study catchments using the PEST tool is illustrated. For the third part, the generation of the different spatial resolutions is first explained, then the uncertainty analysis using both the Generalized Likelihood Uncertainty Estimation GLUE and the Monte Carlo Markov Chain MC^2 methods are explained. In the last part, the implementation of the new concept INPUT is introduced after explaining generation of the different grid aggregation for each catchment.

4.3.1 Setup and Calibration of the WaSiM-ETH Model

4.3.1.1 Choice of the Model Parameters

One of the most important and difficult steps in the calibration and the uncertainty analysis study is the choice of the parameters that are included in both studies. This difficulty arises from the fact that one needs to incorporate as many as possible model parameters. First, to ensure that the interaction between the different parameters are captured and second to guarantee that the model is capable to capture the physical processes occurring at the catchment scale. On the other hand, one has to hold the parameter number as low as possible, in order to radically reduce the computational costs. This should happen not at the cost of simplifying the model at hand.

Thus a sensitivity analysis using one parameter at a time (OAT) was undertaken. Using this method, a large number of parameters are chosen. Then a feasible range for each parameter, either depending on the physical meaning of the parameter, or on literature values, is defined. Within this range the parameter value is changed stepwise, and for each step the model output is calculated. This is done only by changing one parameter value at once. The parameter sensitivity is then calculated using the following equation

$$S = \frac{\Delta O}{\Delta P} \tag{4.8}$$

A large output difference caused by a small parameter change means a large sensitivity of the model to this parameter, and vis versa. Thus, the five most sensitivity parameters have been estimated. They are found in Table 4.6 with their ranges. Similar results are obtained in the study done by Schulla J. (2001), which is found in the WaSiM-ETH handbook.

Table 4.6: The five most sensitive parameters for the WaSiM-ETH model.

Parameter	Description	Unit	Range
m	recession parameter for base flow	$[m]$	$[1 \cdot 10^{-3}, 1]$
T_{kor}	correction factor for the transmissivity of the soil	$[-]$	$[1 \cdot 10^{-4}, 1]$
K_{kor}	correction factor for vertical percolation	$[-]$	$[1, 10]$
SH_{max}	maximum storage capacity of the interflow storage	$[mm]$	$[1, 300]$
KH	single reservoir recession constant for interflow	$[h]$	$[1, 500]$

4.3.1.2 Rainfall Correction

In all WaSiM-ETH runs the rainfall correction module is turned off. This, as explained earlier, is due to the fact that the daily rainfall data are corrected based on the study from Richter (1995). The yearly mean error in the rainfall measurement is taken to be 12% as seen in Figure 4.6. The rainfall data is corrected with a factor 1.12. The corrected rainfall data is then used to simulate the discharge for each catchment.

4.3.1.3 Objective Functions

To evaluate how good the model fits the measured data, a goodness-of-fit criteria is needed. There are different goodness-of-fit measures, in this section two of such measures are explained. For the modeling tool PEST such goodness-of-fit

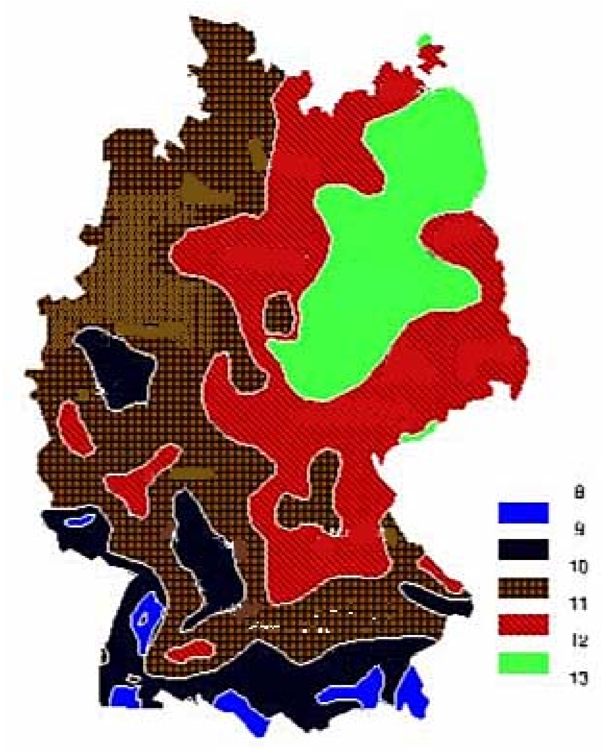


Figure 4.6: Mean correction (%) of the average annual precipitation (1961/90) for moderate Wind-sheltered sites in Germany after Richter (1995).

functions are used as an objective function. The calibration tool PEST uses the sum of squared deviations (ϕ) between measured and simulated data

$$\phi = \sum_{i}^{N} (M_i - O_i)^2 \tag{4.9}$$

whereas, M is the measured value at increment i, O is the observed value at the same increment [Doherty (2004), Moore and Doherty (2005)]. The aim of the calibration tool PEST is to minimize this function, in other words the error between the measured and the simulated data should be minimum. This objective function is sensitive to outliers [Legates and McCabe (1999)] and more weight is added to larger errors.

The second goodness-of-fit and at the same time widely used in hydrological modelling is the Coefficient of Efficiency, or *Nash Sutcliffe* (NS) criteria. It is defined as

$$NS = 1 - \frac{\sum_{i}^{N}(M_i - O_i)^2}{\sum_{i}^{N}(O_i - \hat{O}_i)^2} \tag{4.10}$$

The Nash Sutcliffe criteria represent the mean square error MSE

$$MSE = N^{-1}\sum_{i}^{N}(M_i - O_i)^2 \tag{4.11}$$

divided over the observed data variance subtracted from unity [Legates and McCabe (1999)]. If the MSE of the model prediction is equal to the variance in the observed data, i.e. $NS = 0$, this means that the observed mean is as good as the model predicition, while negative NS values indicate that the observed mean is a better predictor than the model. The Nash Sutcliffe criteria is more sensitive to outliers than the sum of squared errors.

4.3.1.4 Calibration

After choosing the parameters to be calibrated, the PEST files are prepared, for more details the reader is refered to the User Manual [Doherty (2004)]. A warm up period of one year was introduced at the beginning of the calibration period. This step was also made for both MC^2 and GLUE methods. In Table 4.7 the calibration and validation periods for each catchment are seen. To ensure a good match of the simulated data with the measured data, extra parameters, seen in Table 4.8, are calibrated for one time after estimating the first five parameters. After calibrating the extra parameters a last calibration run is made with the five most sensitive parameters. PEST was run using the multipe starting points algorithm within a parallel frame work. This was done to speedup the calibration process.

Table 4.7: The calibration and validatioin periods for all cachments.

Catchment	Calibration	Validation
Weida	Nov. 97 till Oct. 00	Nov. 99 till Oct. 02
Gera-Langenberg	Nov. 96 till Oct. 99	Nov. 93 till Oct. 96
Döllnitz	Nov. 98 till Oct. 01	Nov. 95 till Oct. 98

Table 4.8: Additional parameters considered in the calibration procedure.

Parameter	Description	Unit	Range
P_{grenz}	precipitation intensity threshold for generating preferential flow into the saturated zone	$[mm/h]$	[1,10]
r_k	scaling of the capillary rise/refilling of soil storage from interflow	[-]	[0,1]
c_{melt}	fraction on snowmelt which is surface runoff	[-]	[0,1]
KD	single reservoir recession constant for surface runoff	[h]	[1,72]
T_0	the threshold temperature for snow melt	℃	[-0.5,1.5]
c_0	the degree-day-factor TGF	$[mm/(℃\ d)]$	[0.1,10]

4.3.2 Analysis of Spatial Resolution and Model Parameter Uncertainty

4.3.2.1 Grid Aggregation

To account for uncertainty of spatial resolution, different grid aggregations are generated. This applies for the digital elevation model, the soil map and the land use map. Table 4.9 shows the different grid resolution generated for each catchment. Each grid resolution was created from the original grid resolution.

Table 4.9: The generated grid resolutions from the original resolution for each catchment.

Catchment	Spatial Resolution
Zeulenroda	100, 200, 300 and 500 m
Gera-Langenberg	500, 1000, 2000, 4000 and 6000 m
Doellnitz	100, 300, 500 and 1000 m

For example, to generate the 100 m grid aggregation of the digital elevation model of the Weida catchment, the original 25 m digital elevation model was used. The same procedure is used for the 200, 300 and 500 m grids, respectively. The topographical index was then produced from the corresponding digital elevation model. For creating the different digital elevation model resolutions the cubic convolution method was applied. It determine the new values based on the weighted distance average of the sixteen nearest cells *. For the two other maps the nearest neighbor method was implemented. It determine the location of the

---

*Arcview Help Tutorials

closest cell center in the input grid and assign the value to the cell in the output grid [†].

4.3.2.2 Entropy Theory and Information Content

The Entropy theory from Brasington and Richards (1998) allows the quantification of the information content of continues data sets, in this case both the digital elevation model and the topographical index map. Thus providing a mean of quantifying the lost of information content due to disaggregation. This is helpful when comparing different grid aggragation among each other. The entropy of a system or the information content of a grid is estimated using the following equation

$$I = -\sum_{1}^{\beta} P_i \cdot log(P_i) \tag{4.12}$$

whereas, P is the probability of the values in certain intervals, and β is the number of intervals.

4.3.2.3 Implementing the GLUE Methodology

For GLUE after sampling randomly from the uniform distribution of each parameter, a model run is made. This is repeated 10000 times. For each run the Nash Sutcliffe criteria is estimated. A threshold for behavioral model is set at 0.6 for Zeulenroda and 0.4 for the remaining two catchments. These values are based upon previous experience that has been made with the catchments. The reason for this is to ensure that some models are accepted as behavioral. For the uncertainty bounds, the 90% quantile for the simulated discharge using all behavioral models is used.

4.3.2.4 Implementing the MC^2 Methodology

Again the same parameters ranges and condition, applied to the calibration using PEST, were implemented also here for the uncertainty analysis using Monte Carlo Markov Chain method. In this study only the 5 most sensitive parameters are considered for the study. The other parameters were fixed to their optimal value. The number of MC^2 runs without considering the autocorrelation in the discharge data was set to 10000 runs. The parameter distribution update was done each 200 runs. The burn in period was set to 5000 runs. For the MC^2 with autocorrelation, the total number of runs was 15000 runs, and the burn in period was 10000 runs. After disregarding the burn in period, the results are interpolated. For the parameters a histogram is generated, using the sampled parameter values, with 25 intervals to illustrate the posterior parameter distribution. On the other

---

[†] See Above

hand, for estimating the parameter and the predictive uncertainty bounds, the discharge data are analyzed. For the predictive uncertainty bounds, the discharge for each time step is sorted and the 90% quantile are estimated. For the parameter uncertainty bounds, a normal distribution for each simulated discharge value is generated using the simulated value as mean and the model variance as variance. Having done this and using the generated discharge the 90% quantile are then estimated.

4.3.3 Analysis of Interaction between Rainfall and Model Uncertainty

4.3.3.1 Generating Rainfall Time Series

To represent the rainfall uncertainty in this study, the rainfall measurement error was used. This is bases on the study from Richter (1995). The uncertainty in the rainfall was taken for **point measurements**. This means for the daily rainfall measurement a distribution was generated for each time step, and from this distribution samples were taken for that time step. A normal distribution for the rainfall uncertainty was assumed. This was also found by a study done by Zhihua and Mingelin (2007). The normal distribution for each time step takes the form

$$N(P \cdot (1 + x), P \cdot y) \tag{4.13}$$

where the mean is the daily precipitation P multiplied with the mean error x, taken from the study of Richter (1995). For this study a rainfall measurement error of 12% is assumed, resulting in a mean of $P \cdot 1.12$. The variance σ^2 is assumed to be 5% of the daily values, meaning $\sigma^2 = 0.05$.

4.3.3.2 Implementing the INPUT-Concept

The parameter uncertainty is estimated considering uncertainty in the rainfall input data. In the first part both parameter and rainfall are sampled 10000 times and the model is evaluated and the Nash Sutcliffe is determined. The parameters are sampled from a uniform distribution, and the rainfall as explained in the section above. Using the same principle from GLUE, the same thresholds are defined. In the next step the best parameter set is fixed and only the rainfall is sampled for 1000 times. For the overall uncertainty bound, the 90% quantile are estimated using the results from the first run. For determining the quantiles for the second run, i.e the parameter uncertainty bounds considering input uncertainty, Equations 3.33 till 3.36 are applied.

Chapter 5

Results and Discussion

5.1 WaSiM-ETH Calibration with different Spatial Resolutions

5.1.1 Weida Catchment

Besides the chosen grid aggregations for the uncertainty analysis, an additional grid aggregation, 25 m, was calibrated using PEST. For this grid aggregation, the additional parameters were also considered in the calibration process, see Table 4.8. After calibrating the 25 m grid resolution, for all possible parameters, only the five most sensitive parameters for the other aggregations were calibrated and the additional parameters were fixed at the values found for the 25 m resolution. In Table 5.1 the water balance, the Nash Sutcliffe criteria and the parameter values could be seen. All the parameter values lie in the same order of magnitude. This is also true for the Nash Sutcliffe. The discharge curves for the 25 m aggregation, Figure 5.1 (Figures B.3 to B.5, page 156 to 157 for the rest),

Table 5.1: Water balance, Nash Sutcliffe and parameter values from the calibration of the Weida Catchment using PEST for all grid resolution.

	25 m	50 m	100 m	300 m	500 m
Water Balance (471 mm)	434 mm	491 mm	495 mm	414 mm	393mm
Nash Sutcliffe	0.82	0.78	0.78	0.74	0.81
m	3.07E-2	4.33E-2	4.84E-2	9.87E-3	1.06E-2
T_{kor}	1.00E-4	1.37E-4	4.77E-4	1.21E-4	1.00E-4
K_{kor}	5.42	2.25	2.25	2.25	1
SH_{max}	36.8	41.6	41.9	41.3	39.0
KH	209.5	242.5	246.4	235.1	214.9

show that the calibrated model is able to simulate the dynamic in the catchment. Some of the peaks, seen in Figure 5.1 are well met (e.g. 15.12.99), others are not (e.g. 24.01.00) and some are overestimated (09.03.00). For the time period

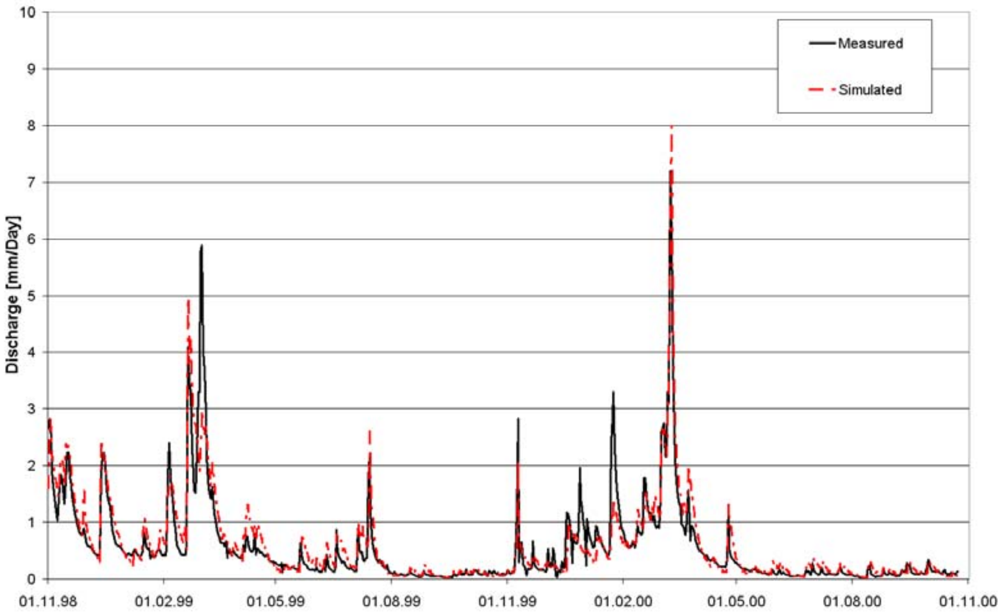


Figure 5.1: The measured and simulated discharge for the Weida catchment using PEST and a 25 m grid resolution.

between December 1999 till mid January 2000 the simulated discharge differs from the measured discharge. This is due to the fact that in this period soil frost occurred, and since the WaSiM-ETH model does not have a routine to deal with soil frost, it is not capable to estimate and simulate the discharge in this time period.

It is clearly seen that the model has the ability to simulate the catchment and its dynamics with a satisfying degree. Thus, an uncertainty analysis could be undertaken to improve the understanding of the system at hand. The uncertainty analysis methods were applied using the same parameter ranges from the calibration method.

5.1.2 Döllnitz Catchment

For the Döllnitz catchment, only the 25 m grid resolution was calibrated. As seen from the calibration of the catchment Weida, all grid resolutions have achieved comparable results. In order to save computational time only the 25 m grid resolution was calibrated for the five most sensitive parameters and the additional parameters. In Table 5.2 the calibrated parameter values and the Nash Sutcliffe are seen. Relatively low Nash Sutcliffe criteria is obtained when compared to

Table 5.2: The Nash Sutcliffe and the parameter values from the calibration of the Döllnitz catchment using PEST for the 25 m grid resolution.

m	T_{kor}	K_{kor}	SH_{max}	KH	Nash Sutcliffe
0.001	1.35E-4	1	143.94	300	0.52

the Weida catchment, though it is still above the acceptance threshold of 0.4. Both parameters m and T_{kor} values are located at the lower boundary of the given parameter range (see Table 4.8 in page 59). This implies that the feasible parameter range should be extended to incorporate parameter values that might lead to a better calibration of the model. Using these calibrated parameter values, a simulated water balance in the catchment of 213 mm is achieved compared to a measured water balance of 196 mm in the two years duration. The hydrograph in Figure 5.2 shows that the major discharge event (17.03.00) is well captured by the WaSiM-Model. The two other major peaks (30.01.00) and (27.03.01) are underestimated and the only summer peak (08.07.01) too.

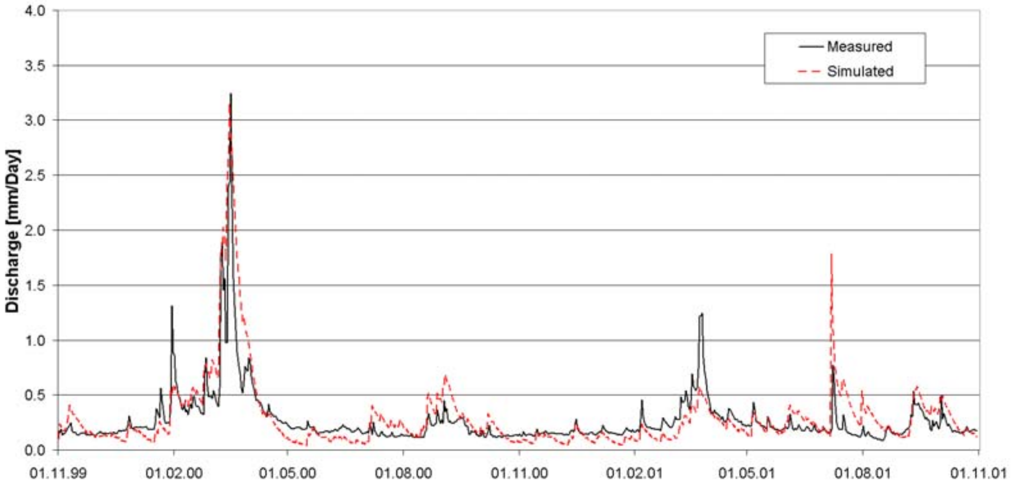


Figure 5.2: The measured and simulated discharge for the Döllnitz catchment using PEST and a 25 m grid resolution.

The basis discharge is not well reproduced by the model. The major peak has a daily discharge of 3.24 mm/day. The base flow is below 0.5 mm/day. Which is a challenge for the WaSiM-model, since the conceptual TOPMODEL was designed for mountainous catchments. Another point is the fact the 8% of the catchment land use is settlements and paved areas, resulting in a fast discharge response. This leads to a spiky hydrograph, as seen in Figure 5.2.

5.1.3 Gera-Langenberg Catchment

For the Gera-Langenberg catchment only the 100 m grid aggregation was calibrated. This was done to test whether the WaSiM-ETH model is capable of representing the catchments in a sufficient way. The same parameters were calibrated, as for the two other catchments. The additional parameters, seen in Table 4.8, are also calibrated once. In Figure 5.3 the simulated discharge is plotted against the measured discharge. It is clearly seen that the WaSiM-ETH model can reproduce the dynamic in the catchment, particularly for the first half of the second hydrological year. The rest of the peaks are not well estimated and the base flow is overestimated for the whole two years. Due to heavy rainfall (over 20 mm/day) the three summer events (13.06.98, 07.07.99 and 15.07.99) are not well captured by the model, indicating the deficiency of the model to simulate fast surface flow due to heavy rainfall.

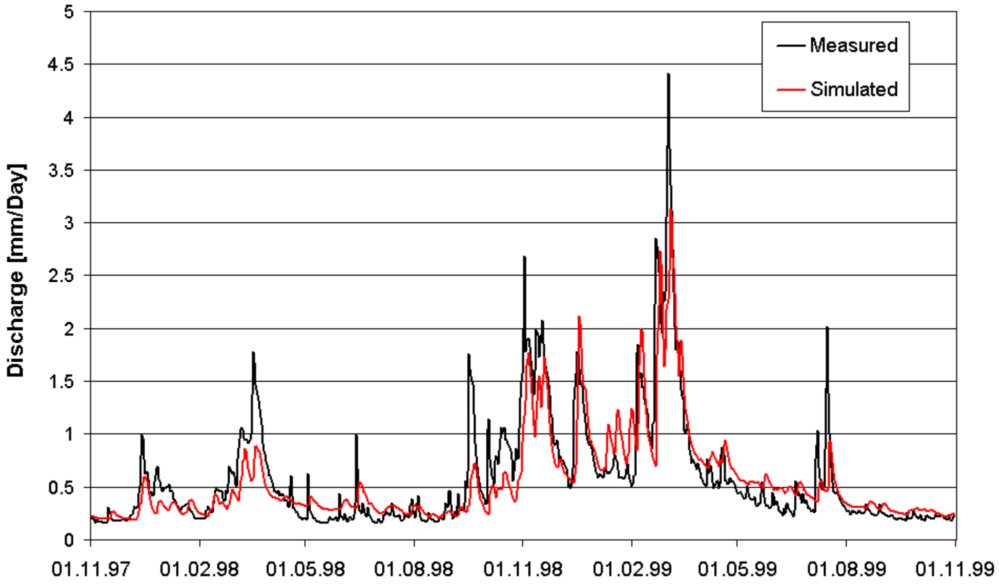


Figure 5.3: The measured and simulated discharge for the Gera-Langenberg catchment using PEST and a 100 m grid resolution.

This could be due to the fact that the WaSiM-ETH model assigns only one parameter value for the whole catchment. For example, the same parameter value for the parameter m is appointed for both the mountainous upper catchment and the plain lower catchment. In Table 5.3 are the calibrated parameter values, the water balance and the Nash Sutcliffe. The water balance is simulated reasonably, less than 10% error, and acceptable Nash Sutcliffe coefficient are achieved. The reason for this is the fact that the Nash Sutcliffe criteria is highly sensitive to outliers, and almost all peaks in the first half of the second hydrological year are

Table 5.3: The water balance, Nash Sutcliffe and the parameter values from the calibration of the Gera-Langenberg catchment using PEST and a 100 m grid resolution.

m	T_{kor}	K_{kor}	SH_{max}	KH	Nash Sutcliffe	Water Balance
6.66E-2	2.08E-4	7.73	247.9	91.2	0.71	408 (430) [mm]

captured.

5.2 Uncertainty Analysis - Spatial Resolution

5.2.1 Small Scale Lower Mountain - Weida Catchment

5.2.1.1 The GLUE Methodology

After calibrating the model and finding no substantial differences among the different grid aggregations, the first uncertainty analysis methodology is implemented. As explained before (see section 3.1, page 25) the aim of this method is to find ***all possible behavioral*** models. Figure 5.4 shows the 95% quantile (black line) and the mean (blue line) for the 200 m grid resolution. These quantiles were generated taking the 95% of all behavioral models. This means the quantile boarders 95% of all behavioral models found. In other words, the wider the range of the quantile, the more uncertain the model is representing the catchment.

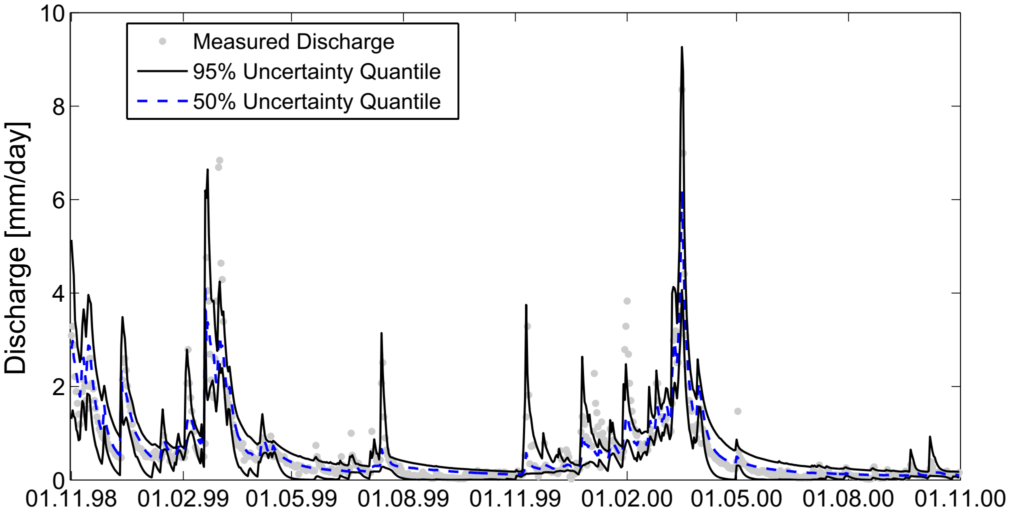


Figure 5.4: The 95% and 50% uncertainty quantile for the Weida catchment using the GLUE method and a 200 m grid resolution.

Studying Figure 5.4, the 95% quantile are close to each other and also, almost all measured discharge points lie within these boundaries, implying a good and a certain representation of the catchment. Comparing the 200 m grid resolution with the 500 m grid resolution Figure B.8 page 158, no obvious differences are seen. This is the same when taking into consideration all grid aggregations (See Figure B.6 till B.8 from page 157 till page 158).

Table 5.4 shows the best parameter values, their Nash-Sutcliffe and water balance values. Similar Nach Sutcliffe values are archived, regardless of the grid aggregation. All parameter values lie within the same magnitude, with exception of the T_{kor} parameter for the 500 m grid resolution. The parameter T_{kor} is associated with topographical index through equation 4.5. In Table 5.5 are the information content of the digital elevation model and the topographical index for the different grid resolutions. The lower value for the parameter T_{kor} could be explained using Equation 4.5. Since the parameter m value is comparable for all grid resolutions, the only different is in the result from the log term, and due to the fact that the topographical index of the 500 m grid resolution differ significantly from the others, see Table 5.5, the value of the parameter T_{kor} should change in order to compensate the topographical index value. The same is true for the water balance, expect for the 100 m grid resolution. In this case an error of almost one fifth the measured discharge is found. Despite this large error in the water balance, comparable values of the Nash Sutcliff 0.79 are estimated. This error arises from the fact that the model over estimates the first month of the calibration period, and in general an overestimation of the smaller peaks are noticed. The high Nash Sutcliffe is due to the fact that the model has managed to simulate the largest peak (17.03.00), these are clearly seen when looking at the hydrograph, Figure B.6 page 157 generated using the best parameter set found for the 100 m grid resolution.

Table 5.4: The water balance, Nash Sutcliffe and parameter values based on the GLUE Method for the Weida catchment for all grid resolution

	500 m	300 m	200 m	100 m
Water Balance (471 mm)	434.3	465.3	437.3	565.4
Nash Sutcliffe	0.78	0.79	0.79	0.79
m	1.09E-2	7.45E-3	1.76E-2	8.73E-3
T_{kor}	1.11E-4	6.06E-2	8.75E-2	3.81E-2
K_{kor}	7.59	9.36	1.14	6.59
SH_{max}	41.2	42.1	44.7	44.4
KH	186.3	206.2	216.5	206.7

Using the best parameter set, the hydrographs for each grid aggregation is generated. Figure 5.5 shows the hydrograph for the 200 m grid aggregation. The

Table 5.5: Information content and relative information content loss for the Weida catchment.

	Digital Elevation Model		Topographical Index	
	Information Content	Relative Loss [%]	Information Content	Relative Loss [%]
25 m	2.1451	0.0	0.7378	0.0
100 m	2.1429	0.10	0.6818	7.60
200 m	2.1295	0.72	0.6543	11.33
300 m	2.1117	1.56	0.6436	12.77
500 m	2.0571	4.10	0.5757	21.97

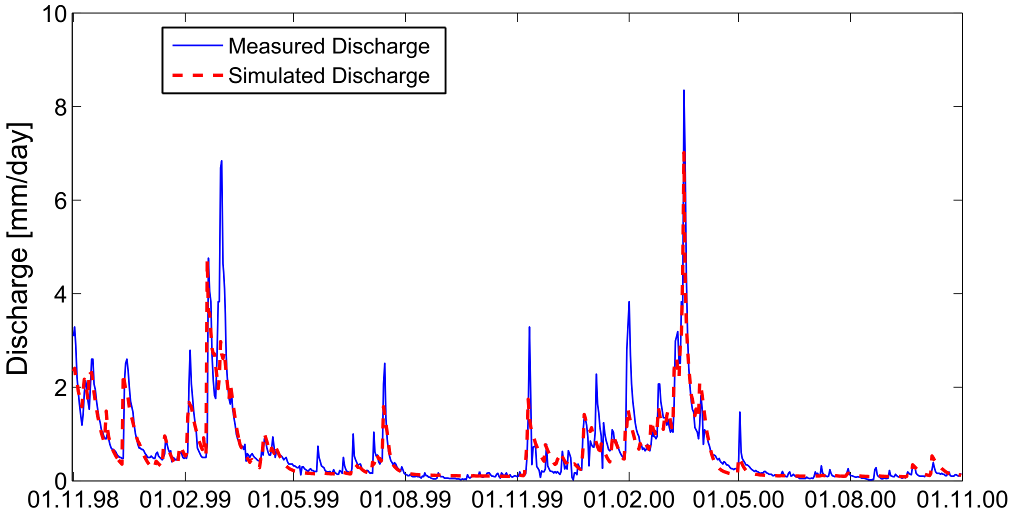


Figure 5.5: The measured and simulated discharge for the Weida catchment using the best paramter set from GLUE and a 200 m grid resolution.

dynamic in the catchment is adequately reproduced. Certain peaks in the hydrograph are captured (e.g. 21.02.99), others are not (e.g. 03.03.99 and 15.07.00). The base flow is also captured with sufficient accuracy. The hydrograph recession is also represented with satisfactory quality. This behavior is also seen for the other grid resolutions (Figures B.9 till B.11 from page 158 till page 159).

Another characteristical result from the GLUE methodology is the dotty plots, seen as an example, for the 200 m grid aggregation in Figure 5.6. It is distinctly clear that for the parameters T_{kor}, K_{kor}, SH_{max} and KH a good Nash Sutcliffe value could be achieved within the whole feasible parameter range. Except for the parameter m, where good Nash Sutcliffe values are only attained within a narrow range of the parameter. This implies that the parameter m is the most sensitive parameter. Similar results are obtained for the other grid resolutions.

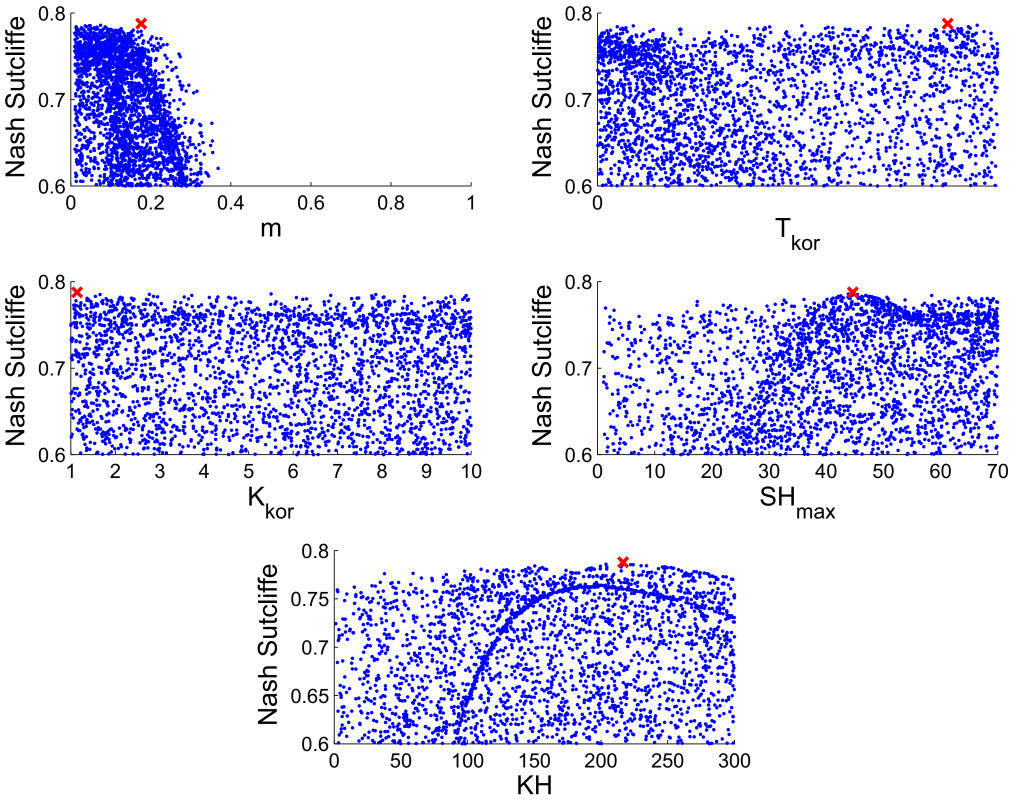


Figure 5.6: The dotty plots for the Weida catchment using a 200 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

For the validation period, good results (for both Nash Sutcliffe and the water balance) are obtained, see Tabel 5.6. The Nash Sutcliffe values lie on the lower end of the acceptable range (>0.6) but still they are within this range. The hydrographs, seen in Figures B.12 till B.14 in pages 159 till 160, show that the dynamic in the catchment is well represented, the peaks are met temporally but the four major events are underestimated. Also in the validation period no significant differences between the several aggregations were found.

Table 5.6: The water balance and the Nash Sutcliffe criteria for the Weida catchment using the GLUE method for the validation period.

Grid Resolution [m]	Water Balance (351 mm)	Nash Sutcliffe
500 m	358.6	0.69
300 m	405.8	0.63
200 m	371.1	0.63
100 m	364.5	0.69

5.2.1.2 The MC^2 Methodology

On the contrary to GLUE, the Monte Carlo Markov chain MC^2 could be considered as a redundant calibration tool, which searches for the best parameter set using a probability framework. This is the reason why the method has the ability to quantify both parameter and predictive uncertainty. This section is split into two parts, the first part handles the case of MC^2 without taking autocorrelation in the measured discharge data into account, and the second part discusses the case of MC^2 with autocorrelation.

MC^2 without Autocorrelation:

The uncertainty bounds, in Figure 5.7 estimated from the MC^2 again for the 200 m grid resolution, one could clearly see that the predictive bounds (black line) are relatively narrow, but for the major peak (03.03.00) the uncertainty bounds are wide compared with GLUE (MC^2 ~ 15 mm/Day and GLUE ~ 9 mm/Day). The parameter uncertainty bounds are almost a single line, implying a negligible role of parameter uncertainty in the system. This also applies for the rest of the grid aggregation, see Figures B.18 till B.20 from page 162 till 163.

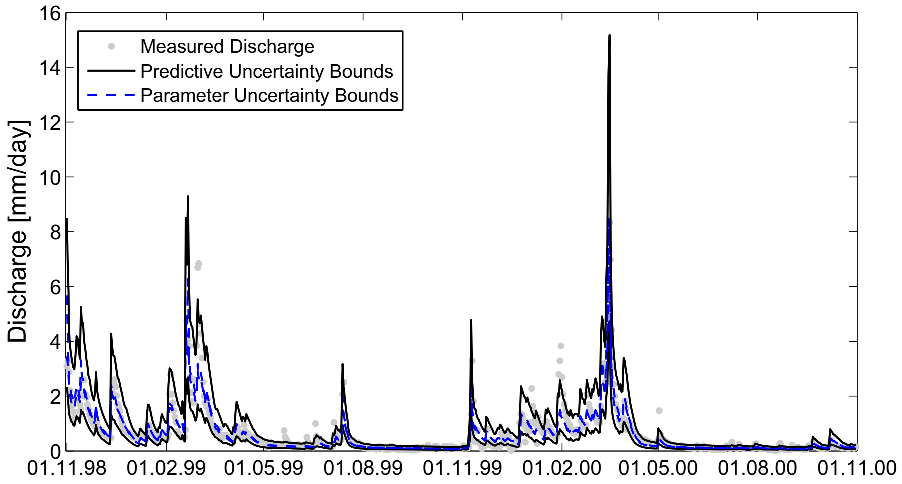


Figure 5.7: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 without autocorrelation and a 200 m grid resolution.

Table 5.7 shows the best parameter set that have been found for each aggregation. As explained in Section 3.2.4 page 33, the model variance parameter serves to separate the parameter uncertainty from the predictive uncertainty. The model variance has a low value. This is the reason for the tight parameter uncertainty bounds.

The small model variance could be explained by looking at the definition of the

Table 5.7: The water balance, the Nash Sutcliffe and the parameter values for the Weida catchment.

	500 m	300 m	200 m	100 m
Water Balance (471 mm)	636	502	495	491
Nash Sutcliffe	0.62	0.72	0.72	0.68
m	1.11E-2	8.76E-3	8.79E-3	1.54E-2
T_{kor}	0.781	0.950	0.565	6.53E-3
K_{kor}	1.67	6.50	9.69	1.12
SH_{max}	39.2	36.1	34.3	37.72
KH	252.1	253.5	241.7	277.1
Model Variance	0.215	0.21	0.207	.147

likelihood function equation 3.14. It is noticed that the non-exponential part is raised to the power $(n/2)$, where n is the number of data points, in this case 730. This results in stretching the parameter space, thus in a sharp peak. The model variance is taken on the peak, and since the peak is sharp, this produces a small model variance. In other words, the definition of the likelihood forces the parameter uncertainty to be small, and the remaining of the uncertainty is considered as the predictive uncertainty of the model.

The discharge curve for the 200 m grid aggregation in Figure 5.8, shows that the dynamic in the system is well represented, and almost all peaks are also met. This could also be seen in the water balance of the system. On the other hand, the 500 m grid aggregation has a large error in the water balance and this could

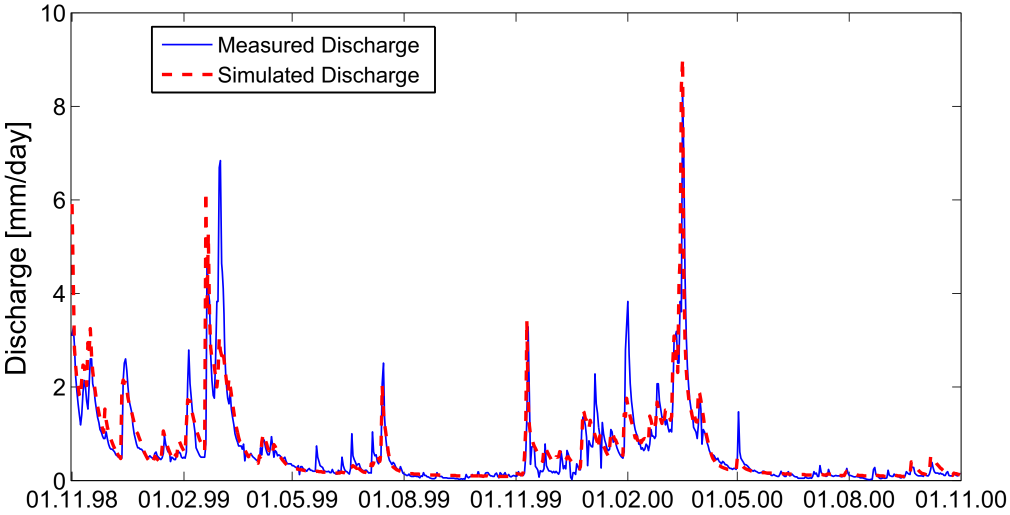


Figure 5.8: The measured and simulated discharge for the Weida catchment using MC^2 without autocorrelation and a 200 m grid resolution.

be clearly seen in the discharge curve, Figure B.23 page 164. Here, the simulated discharge is higher than the measured discharge. The reason for this large error in the water balance for the 500 m grid resolution can be explained by the posterior parameter distribution, Figure 5.9. The parameters have not converged, this is indicated through the wide spread of the parameter over the whole parameter space, and the presence of outlier. Another picture is seen when looking at the posterior parameter distribution for the the 200 m grid aggregation, Figure B.29 in page 166. A clear and well defined distribution is seen for all of the five parameters with the exception of K_{kor}. This suggests that this parameter can not be well defined.

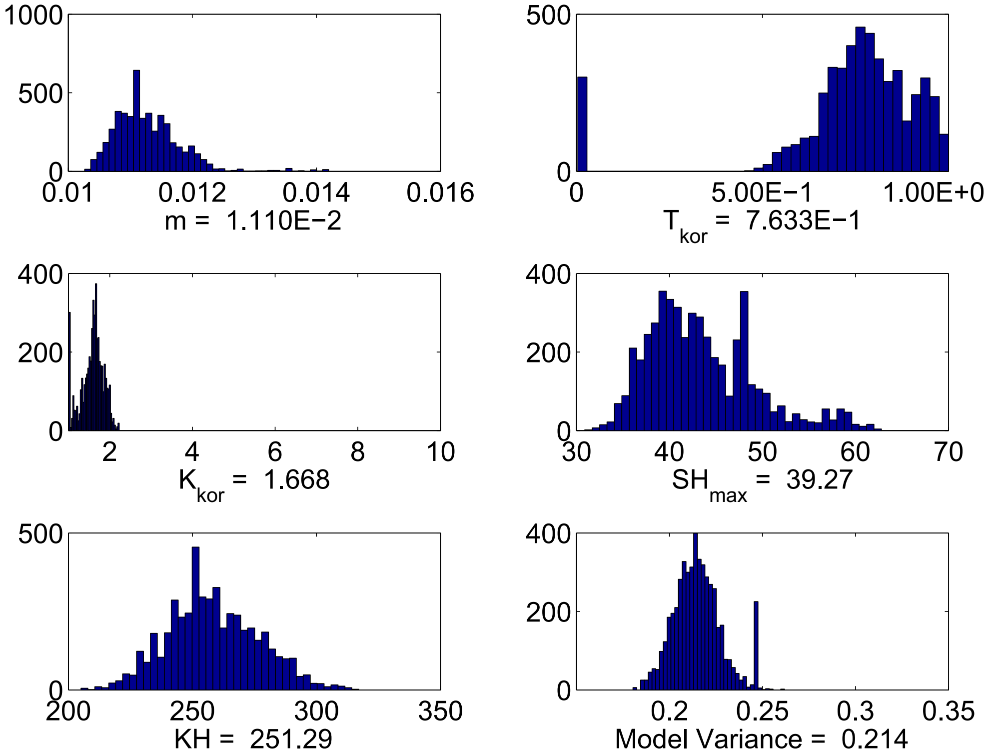


Figure 5.9: The parameter posterior distribution for the Weida catchment using the MC^2 method without autocorrelation and the 500 m grid resolution.

The MC^2 without autocorrelation method did not do well in the validation period, see Table 5.8 but this does not imply that the method is incapable of finding the best parameter set. It is only that this parameter combination does not do well in the validation period. The negative Nash Sutcliffe values indicate that the prediction using the measured discharge mean is better than using the model. This is also clearly seen in hydrographs, Figures B.24 to B.27 at pages 164 to 165. The dynamic in the system is badly represented, the timing too and no peaks are met. Comparing Table 5.7 with Table 5.8 only the difference in the

Table 5.8: The water balance and the Nash Sutcliffe criteria for the Weida catchment using the MC^2 without autocorrelation Method for the validation period.

	Water Balance (351 mm)	Nash Sutcliffe
500	457.3	-0.676
300	601.5	-2.45
200	592.4	-2.46
100	636.13	0.646

value of the parameter T_{kor} is noticed. This is the reason for the bad results in the validation period. Because of this value the model might be able to capture some of the dynamic and the physical processes occouring at the catchment scale in the calibration period but not necessarily at the validation period.

MC^2 with Autocorrelation:
A first order autocorrelation is implemented. The autocorrelation assumes that the error in the measured discharge at time (t) contains error from the previous time steps. For a first order autocorrelation, the error at time step (t) is only dependent on the error from the previous time step ($t-1$). In Table 5.9 the water balance, the Nash Sutcliffe, the parameter values, the model variation and the first order autocorrelation coefficient AR are presented. The model variance is lower than the model variance for the case without the autocorrelation coefficient. Implying that the error in the discharge data influenced the model variance, thus indicating a strong interaction between input uncertainty and model uncertainty. The autocorrelation factor AR indicates the percentage of error in time step (t) that can be explained or are due to the error from former time step ($t-1$). Except for the 100 m grid, almost 70% of errors in discharge data can be explained by

Table 5.9: The water balance, the Nash Sutcliffe and parameter values for the Weida catchment using the MC^2 method with autocorrelation.

	500 m	300 m	200 m	100 m
Water Balance (471 mm)	520.8	485.8	474.9	485.6
Nash Sutcliffe	0.738	0.805	0.811	0.82
m	9.387E-03	1.039E-02	9.258E-03	2.002E-02
T_{kor}	0.315106	0.523185	0.74173	2.00E-04
K_{kor}	9.55	8.76	4.90	5.84
SH_{max}	25.0	61.1	57.6	56.3
KH	173.0	356.6	338.0	356.9
Var	0.107	0.118	0.120	0.109
AR	0.70	0.69	0.68	0.58

errors from the time step before, i.e. the day before. For such small catchments with short response time, not all errors in the time step (t) could be explained or are due to the error from the time step ($t-1$). The uncertainty bounds for the case with autocorrelation, Figure 5.10, are narrow, this is true also for the parameter bounds. For the largest peak (03.03.99) the predictive uncertainty bounds are smaller than for the case without autocorrelation (without 16 mm/day, with 10 mm/day) , thus indicating a role of the error in the discharge data and it's contribution to the predictive uncertainty of the model.

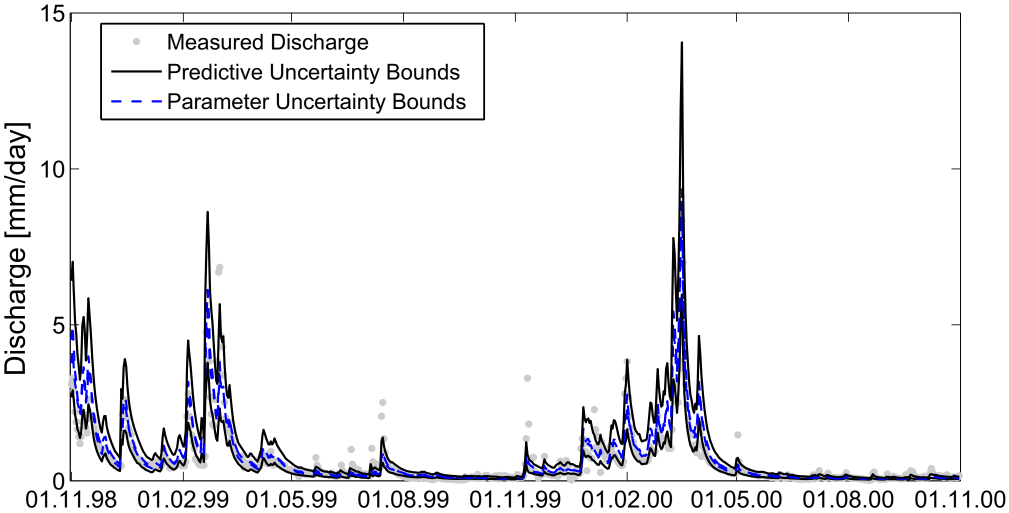


Figure 5.10: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 with autocorrelation and a 100 m grid resolution.

The posterior parameter distribution, for the 100 m grid resolution seen in Figure 5.11, the posterior parameter distributions for the rest of the gird resolutions are seen in Figures B.34 till B.36 page 169 till 170) indicates a good convergence of the parameter m, model variance and the autocorrelation coefficient AR. The simulated discharge curves in Figure 5.12 show almost a prefect fit to the measured one. This is again due to the error elimination with the autocorrelation factor.

For the MC^2 with autocorrelation in the validation period, good results are also obtained, Table 5.10. Though one point should be mentioned, the MC^2 with autocorrelation could be implemented, since discharge data is available for the validation period. This is a disadvantage of the method, since it could not be used to extrapolate and simulate the water balance in the catchment due to change in management changes in the catchment. The discharge curves for the validation are seen in Figures B.40 till B.43 in 172 till 173. The discharge curves mirror the water balance results and the Nash Sutcliffe seen in Table 5.10

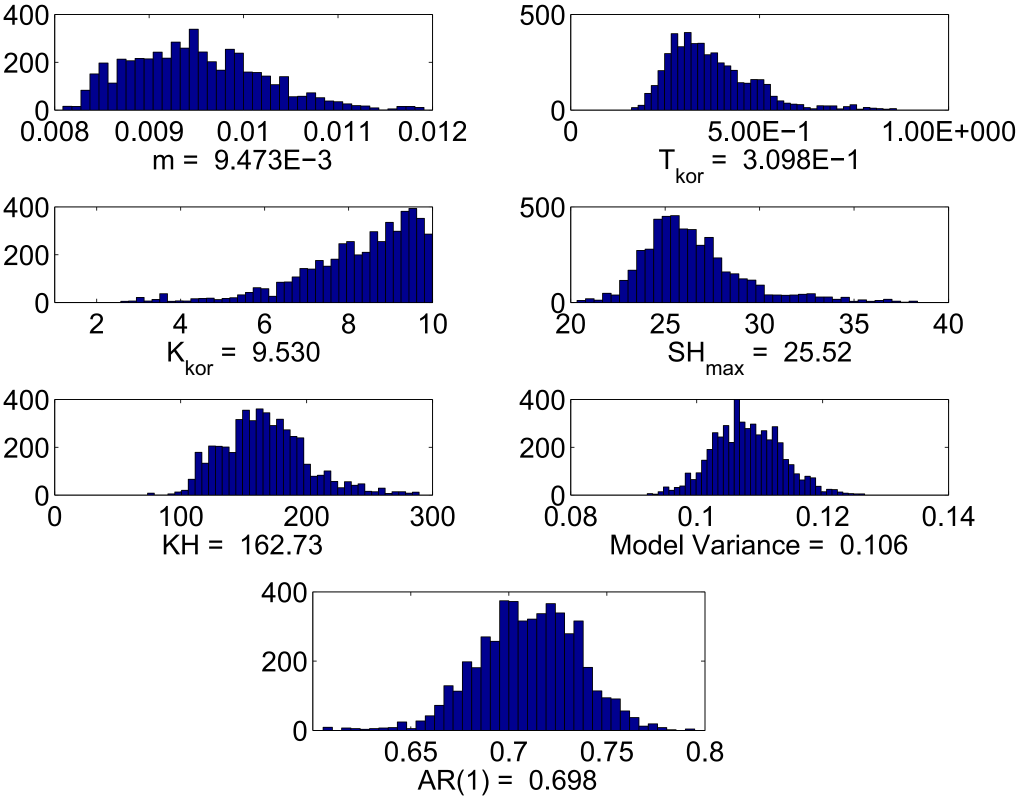


Figure 5.11: The parameter posterior distribution for the Weida catchment using the MC^2 method with autocorrelation and the 100 m grid resolution.

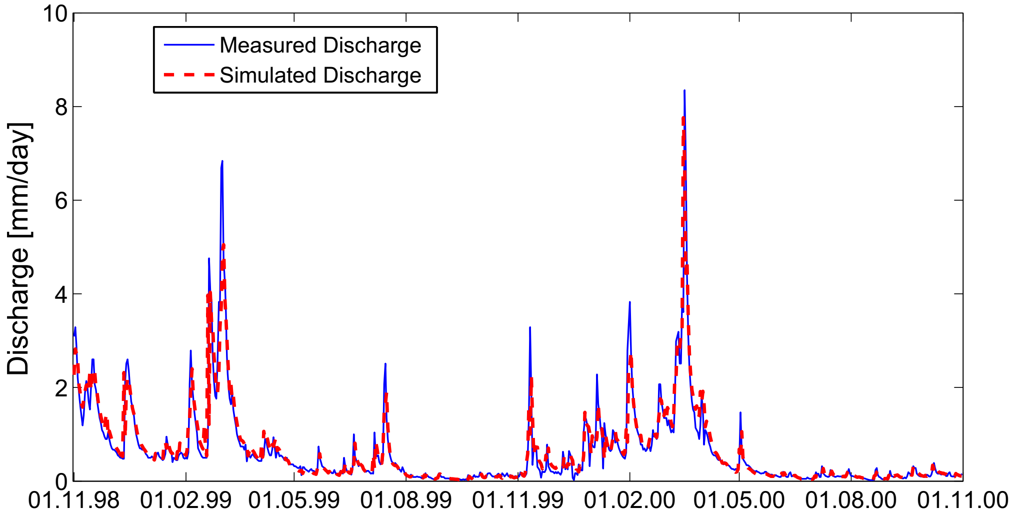


Figure 5.12: The measured and simulated discharge for the Weida catchment using the MC^2 with autocorrelation and 100 m grid resolution.

Table 5.10: The Water balance and the Nash Sutcliffe criteria for Weida using MC^2 with autocorrelation for the validation period.

	Water Balance (351 mm)	Nash Sutcliffe
500	455	-2.3
300	603	-0.95
200	577	-1.4
100	363	0.635

5.2.2 Small Scale Loess Catchment - Döllnitz Catchment

5.2.2.1 The GLUE Methodology

As explained in the previous section the parameter feasible ranges were extended to ensure a large number of behavioral models. Table 5.11 shows the parameters values, Nash Sutcliffe coefficient and the water balance for the different grid aggregations. By extending the feasible parameter range, better results are obtained using the GLUE methodology. For the parameter m the values lie above the lower bound of the parameter range, but the parameter KH values are well above the initial parameter range. For all grid resolutions, except the 1000 m, comparable Nash Sutcliffe value have been obtained. This could be explained by the information content of each grid resolution in Table 5.12. The information content lost due to the aggregation to 1000 m is much larger than due to aggregation to 500 m. Due to this information content loss, a lower Nash Sutcliffe criteria is achieved by the 1000 m.

Table 5.11: The water balance, the Nash Sutcliffe criteria and the parameter values for the Döllnitz catchment using GLUE.

	1000	500	300	100
Water Balance (196.5 mm)	224.5	237.1	255.4	237.4
Nash Sutcliffe	0.54	0.59	0.60	0.60
m	3.153E-2	9.278E-3	7.447E-3	1.183E-2
T_{kor}	4.176E-3	8.019E-1	6.062E-2	2.352E-1
K_{kor}	12.1	1.5	9.4	2.0
SH_{max}	91.7	82.6	42.1	82.2
KH	997.8	873.1	206.2	843.1

The dotty plot of the behavioral models (Figure 5.13) for the 300 m grid resolution as an example shows that (again) m is the controlling parameter, where behavioral models only exists within a small range between 0.001 and 0.05. For all other parameters, a good Nash Sutcliffe value is obtained within the whole

Table 5.12: Information content and information content loss for the different grid resolution for the Döllnitz catchment.

	Digital Elevation Model		Topographical Index	
	Information Content	Error	Information Content	Error
25	2.1481		0.74571	
100	2.1468	0.06	0.7075	5.12
300	2.1321	0.74	0.6752	9.45
500	2.0941	2.51	0.6200	16.86
1000	1.9891	7.40	0.4983	33.18

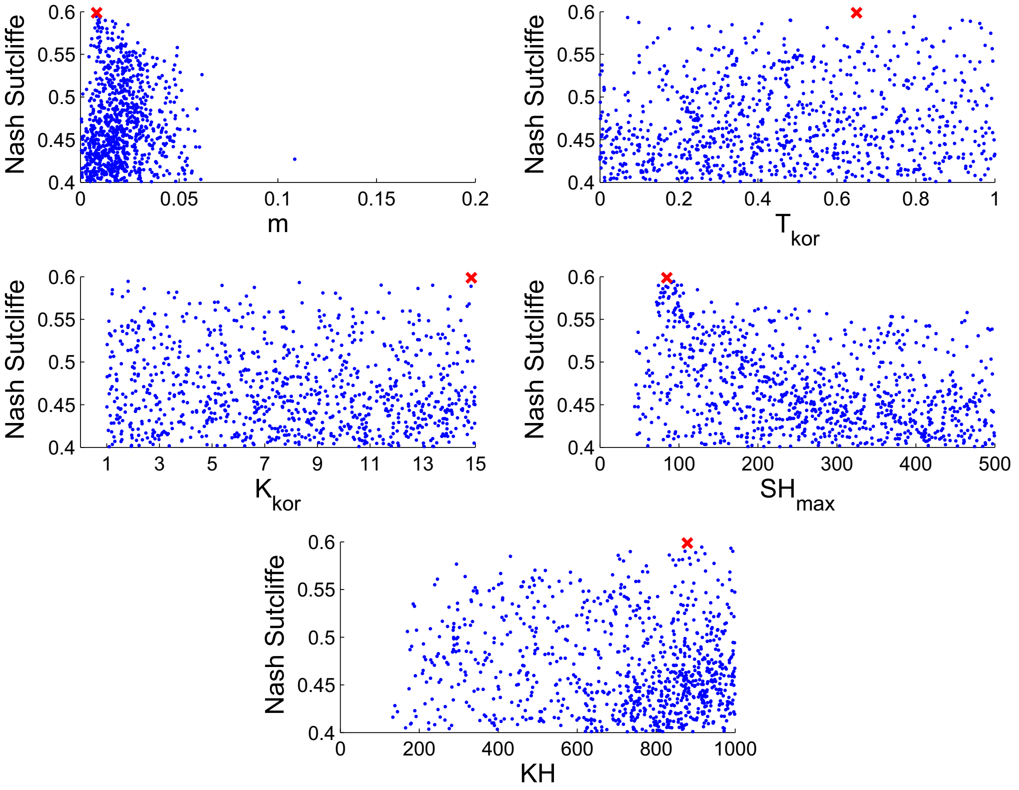


Figure 5.13: The dotty plots for the Döllnitz catchment using a 300 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe

parameter range. This is also true for the other grid resolutions, see Figures C.2 till C.4 from page 182 to 183. The resulting uncertainty bounds from the GLUE uncertainty analysis are shown for the 300 m grid resolution in Figure 5.14. Only half of the daily discharge point (grey) lie within the uncertainty quantiles emphasizing the fact that the WaSiM-ETH model has problems with simulating the base flow discharge. The catchment dynamic is not well captured by the model. Comparing the different grid aggregations, the uncertainty quantiles look almost the same, see Figures C.5 to C.7 from page 183 to 184. For major discharge events, the uncertainty quantiles are wider than for the low flow. Only for the event in summer 2001 (08.07.01) the uncertainty quantiles are identical.

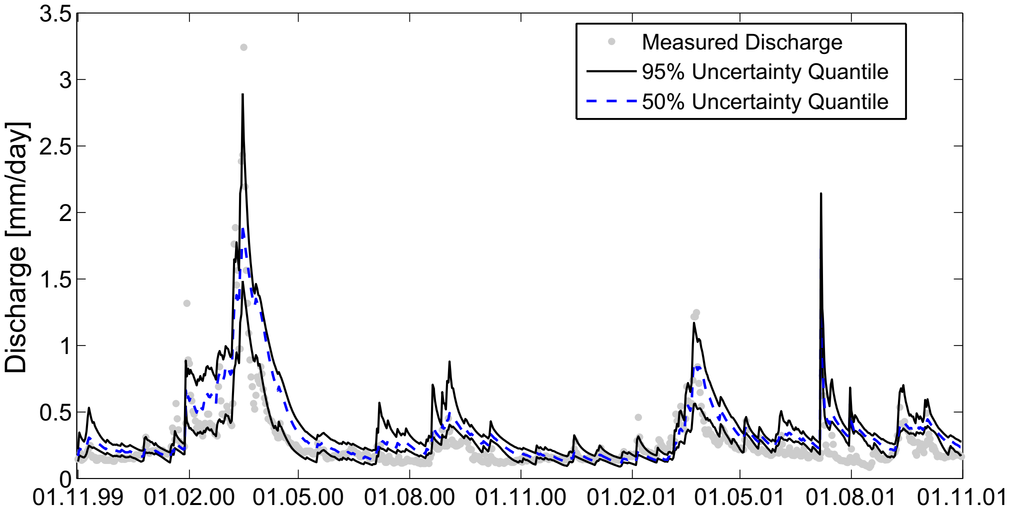


Figure 5.14: The 95% and 50% uncertainty quantile for the Döllnitz catchment using the GLUE method and a 300 m grid resolution.

The discharge hydrograph calculated using the best parameter set found by the GLUE methodology in Figure 5.15 for the 100 m grid resolution, we see that the WaSiM-ETH model, using the 100 m grid aggregation, does not reproduce the dynamics of the system. Nevertheless, the major peaks (discharge > 0.5 mm/day) are simulated better but not well. For the summer event (08.07.01) the discharge is quite overestimated, for both spring events (17.03.00 and 26.03.01), the model overestimates the first event but underestimate the second event. The same picture is seen for the 500 m grid resolution (Figure C.9). A different picture is seen when observing the discharge hydrograph from the 300 m grid resolution in Figure C.8. The hydrograph is spiky. The catchment dynamic is finely or well reproduced and the first winter event is overestimated. For the 1000 m grid resolution, Figure C.10, both winter events are underestimated. Noticeable is that the summer event (08.07.01) is overestimated for all grid aggregations. This indicates a systematic error in the model or an error in the input data. The error

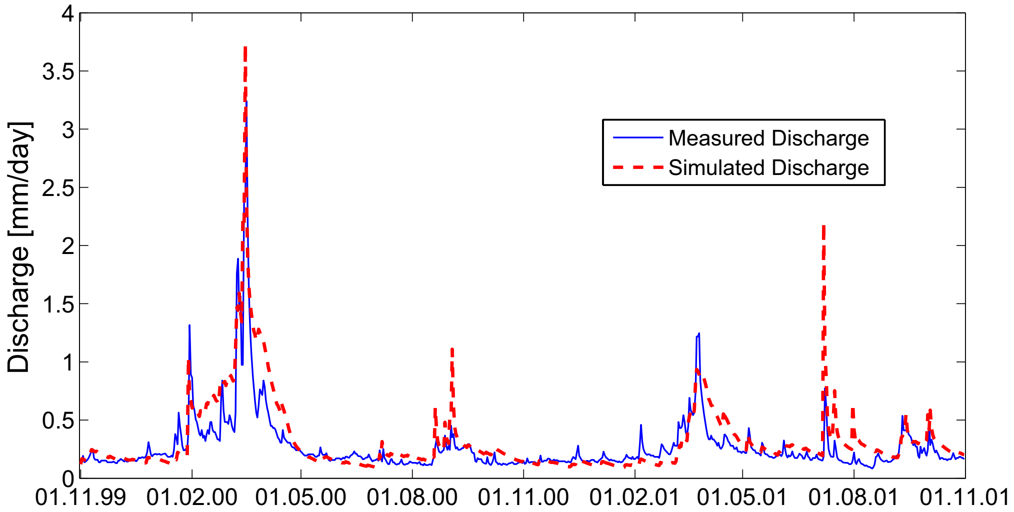


Figure 5.15: The measured and simulated discharge for the Döllnitz catchment using GLUE and a 100 m grid resolution.

in the land use input data could be in the form of inadequate representation of the catchment and its characteristics (for example the paved areas) or in the form of incorrect parameterization of the different land uses. A systematic error in the WaSiM-ETH model could be a missing routine in the model which could handle direct discharge and the routing from paved areas. This theory is confirmed by looking at the uncertainty quantiles for the different grid aggregations (Figure 5.14 and Figures C.5 to C.7 from page 183 to 184). The uncertainty quantiles for this event lie almost on each other implying that the model is not capable of catching the physical process behind the formation of surface runoff.

The spiky hydrograph for the 300 m grid aggregation is due to the value of the parameter SH_{max}, see Table 5.11. The value of this parameter is half the value of the same parameter for other grid aggregations. This parameter controls the maximum storage of the interflow storage. When the interflow storage is filled, surface discharge is generated. Since for the 300 m grid resolution the parameter SH_{max} is half as large as the other grid resolutions, it is filled much faster and the rest of the water is used for surface runoff generation, thus explaining the spiky discharge.

In the validation period the GLUE method does achieve reasonable results, see Table 5.13. The Nash Sutcliffe values are negative, the water balance has an error more than 25%. As explained earlier the WaSiM-ETH model with the TOPMODEL approach is not capable of capturing the physical processes behind the generation of discharge in this catchment.The discharge curves for the validation period are seen in Figures C.12 to C.14 in pages 186 to 187

Table 5.13: The water balance and the Nash Sutcliffe Criteria for Döllnitz catchment using the GLUE method for the validation period.

	Water Balance (192.6 mm)	Nash Sutcliffe
1000	229.4	-0.07
500	251.9	-2.9
300	260.6	-0.73
100	250.6	-0.28

5.2.2.2 The MC^2 Methodology

MC^2 without Autocorrelation:

Table 5.14 shows the water balance, the Nash Sutcliffe criteria and the parameter values. Noticeable is the change of the order of magnitude of 3 for the parameter (m) as well as for (T_{kor}) for the 100 m grid resolution. This is due to the fact that the parameters (m) and (T_{kor}) did not converge, as seen in the posterior parameter distribution for the 100 m grid resolution in Figure 5.16. This also explains the relatively larger model variance obtained for the 100 m aggregation. The posterior parameter distribution for the other grid resolutions are seen in Figures C.15 to C.17 from page 188 to 189, all parameters have converged.

Table 5.14: The water balance, Nash Sutcliffe criteria and the parameter values, for the Döllnitz Catchment using the MC^2 Method without autocorrelation.

	1000	500	300	100
Water Balance (196.5 mm)	156.1	173.0	170.4	173.5
Nash Sutcliffe	0.294	0.387	0.385	0.39
m	0.0693	0.1363	0.138	8.631E-4
T_{kor}	4.726E-3	2.111E-4	7.5E-5	5.398E-1
K_{kor}	14.3	4.4	4.56	8.0
SH_{max}	46.1	66.6	65.3	138.9
KH	636.6	932.2	977.75	982.5
Model Variance	0.137	0.136	0.133	0.199

In general, the Nash Sutcliffe values are low, below the threshold (>0.4) and the water balance is not simulated with satisfaction, similar results were also obtain using the GLUE methodology. This finding might indicate the inability of the WaSiM-ETH model to simulate the discharge output from the Döllnitz catchment. One reason could be the fact that the TOPMODEL-approach is designed for mountainous steep catchments, whereas the döllnitz catchment is characterized by flat slope and Loess soil. Another point leading to the worse simulation of the catchment, is that the catchment is dominated by groundwater flow and

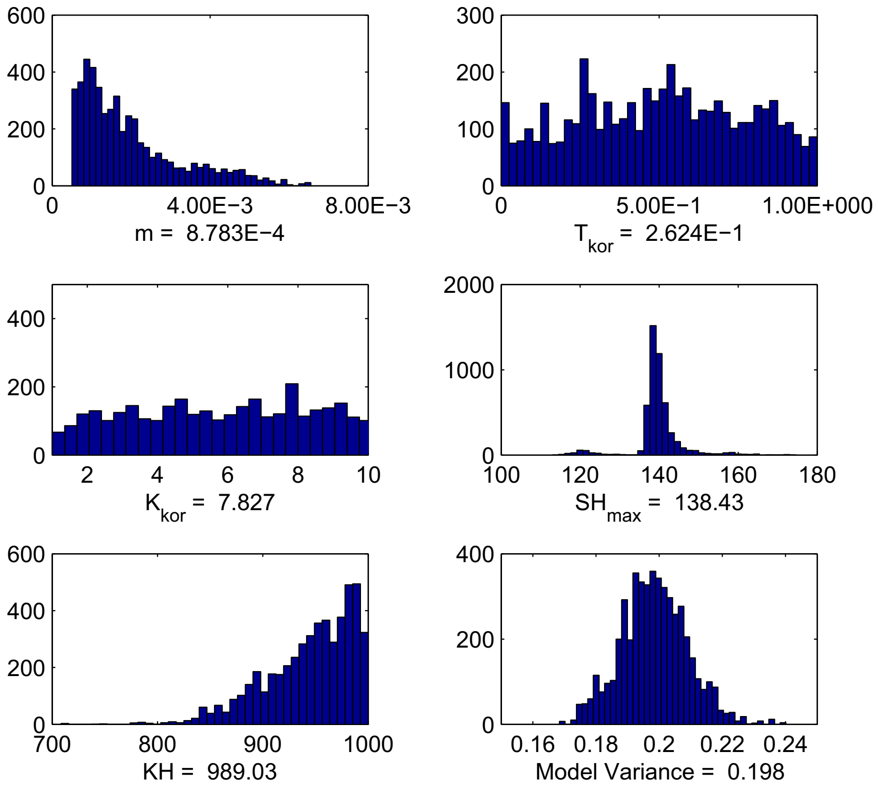


Figure 5.16: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method without autocorrelation and the 100 m grid resolution.

no interflow is found. It is well known that the TOPMODEL-approach is a simplified concept to represent water movement in soils. This implies that the TOPMODEL-approach is not suitable or not capable to represent the catchment in a good manner. To check this theory, one can apply or implement another concept or approach for the soil simulation, e.g. Richard's Equation.

Figure 5.17 shows the simulated discharge for the 300 m grid aggregation. Obviously the model tries to capture the dynamic in the system, with some success. The peak discharges are not met at all, the baseflow too. This supports the theory, that the TOPMODEL-approach does not have the ability to simulate the catchment, because of it's characteristics. One event (08.07.01) the peak is quite overestimated. From the flow components it is clear that this large peak is due to Horton flow in the catchment, coming from the paved areas. For the 100 m grid resolution a different picture is seen. The peaks are poorly met, the dynamic in the system is hardly reproduced, which is a consequence of the un-converged parameters.

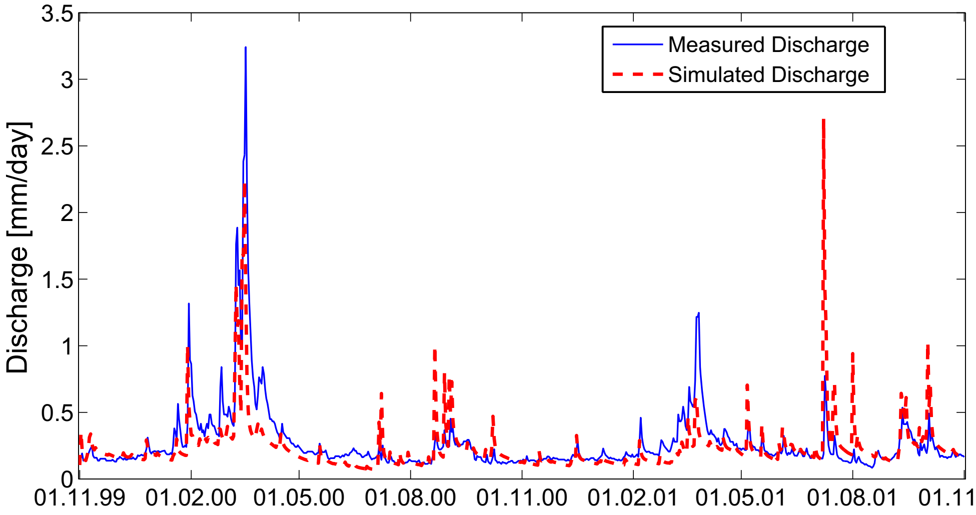


Figure 5.17: The measured and simulated discharge for the Döllnitz catchment using the MC^2 method without autocorrelation and a 300 m grid resolution.

The predictive uncertainty bounds in Figure 5.18 for the 500 m grid resolution are relatively close to each other, the parameter uncertainty bounds are almost one line, due to the small model variance value. Similar results are obtained for the other grid resolutions, except for the 100 m, Figure C.18 in page 190, since the parameter did not converge. In the validation period the parameter

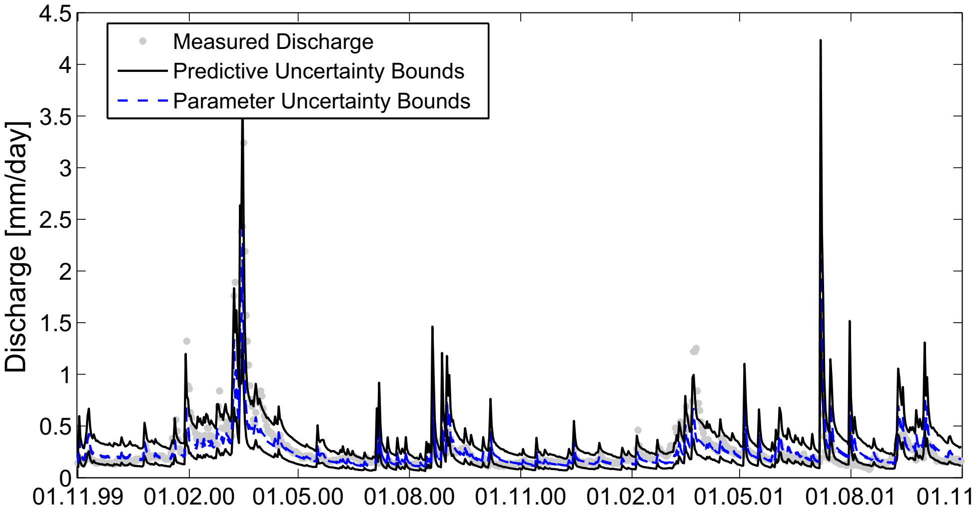


Figure 5.18: The predictive and parameter uncertainty bounds for the Döllnitz catchment using MC^2 without autocorrelation and a 500 m grid resolution.

sets found using the MC^2 without autocorrelation method behave the same the parameter found using the GLUE method. Table 5.15 shows the results for the validation period. Negative Nash Sutcliffe indicates a bad simulation and a the fact that the peak discharges are not met at all. This is despite the fact that the water balance has an error below 10%. The discharge curves for the validation period could be seen in Figures C.21 till C.24 at pages 191 till 192.

Table 5.15: The water balance and the Nash Sutcliffe criteria for Döllnitz using the MC^2 method without autocorrelation for the validation period.

	Water Balance (192.6 mm)	Nash Sutcliffe
1000	173.1	-0.52
500	182.5	-0.34
300	179.0	-0.38
100	179.5	0.22

MC^2 with Autocorrelation:

The hydrograph diagram in Figure 5.19 shows a totally different picture for the 500 m grid resolution. The simulated discharge matches the measured discharge, except for the summer event 08.07.01, the same behavior is also observed here. Table 5.16 shows the parameter values, the Nash Sutcliffe criteria and the water balance. Since almost a prefect fit of the simulated to the measured discharge is achieved, a very good water balance is attained with an error of almost 1%. Despite this, the Nash Sutcliffe criteria is relatively low for such good results (between 0.52 and 0.71). The reason for this is the summer event (08.07.01),

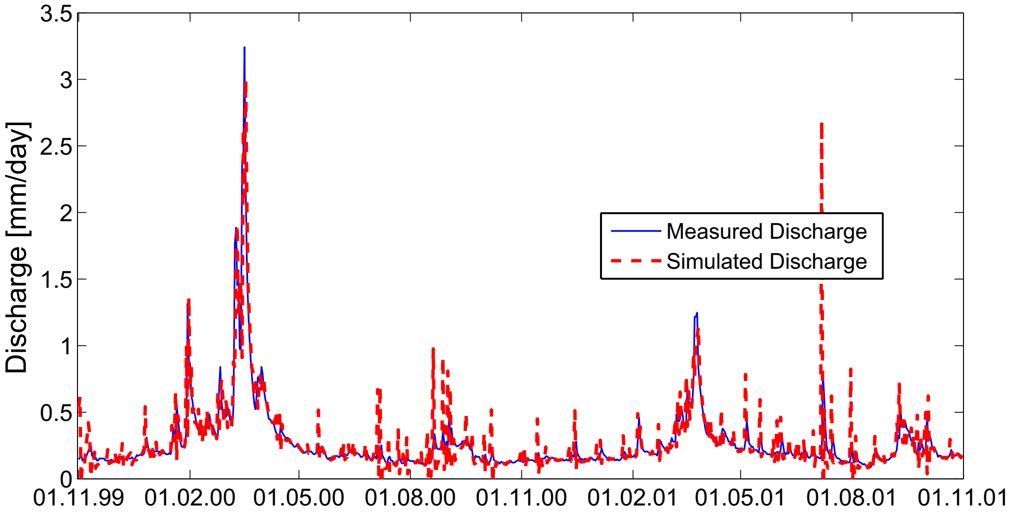


Figure 5.19: The measured and simulated discharge for the Döllnitz catchment using MC^2 with autocorrelation and a 500 m grid resolution.

since the criteria is sensitive to outliers. Recalculating the Nash Sutcliffe criteria disregarding the summer event, values between (0.7 and 0.8) are obtained.

The most important parameter values is that of the first order autocorrelation AR. Almost 99.5% of the error in the discharge can be explained or comes for the time step (i.e. day) before. But not only this, due to the introduction of this extra statistical parameter, the same WaSiM-ETH model is able to simulate and reproduce a correct hydrograph, water balance and the dynamic in the system. This indicates that the discharge in this catchment a major error source is, and if not handled explicitly no good simulations will be obtained.

Table 5.16: The water balance, Nash Sutcliffe Criteria and the parameter values for the Döllnitz catchment using the MC^2 with autocorrelation method.

	1000	500	300	100
Water Balance (196.5 mm)	198.3	198.2	199.0	198.1
Nash Sutcliffe	0.71	0.52	0.67	0.68
m	0.996	0.995	0.996	0.986
T_{kor}	1.223E-01	9.656E-03	2.551E-03	2.876E-04
K_{kor}	14.2	12.5	13.7	12.5
SH_{max}	51.2	418.3	377.6	417.6
KH	88.5	103.0	118.7	106.0
Model Variance	0.032	0.032	0.032	0.032
AR-Factor	0.996	0.994	0.994	0.994

Also a different picture is obtained by the uncertainty bounds. Figure 5.20 shows the uncertainty bounds for the 300 $4m$ grid resolution. Both uncertainty bounds, the predictive as well as the parameter uncertainty bounds, lie on the same line. Indicating a well converged model, and a small predictive uncertainty. Also this finding supports the theory that the discharge a major error source is in this catchment. The posterior parameter distribution does shows the expected picture, all parameter have converged, this is seen in Figure 5.21 for the 300 m grid resolution. In Figures C.25 to C.27 from page 193 to 194.

The introduction of the autocorrelation factor has also helped in achieving good results for the catchment. Table 5.17 show the results obtained. This is also clear in the hydrographs, seen in Figures C.28 to C.31 from pages 194 to 195. Though, the MC^2 with autocorrelation was able to attain good results, this was only possible since measured discharge data is available for the validation period. For forecasting future event is this method not suitable, since no measurement data exists to be able to estimate the error at any time step.

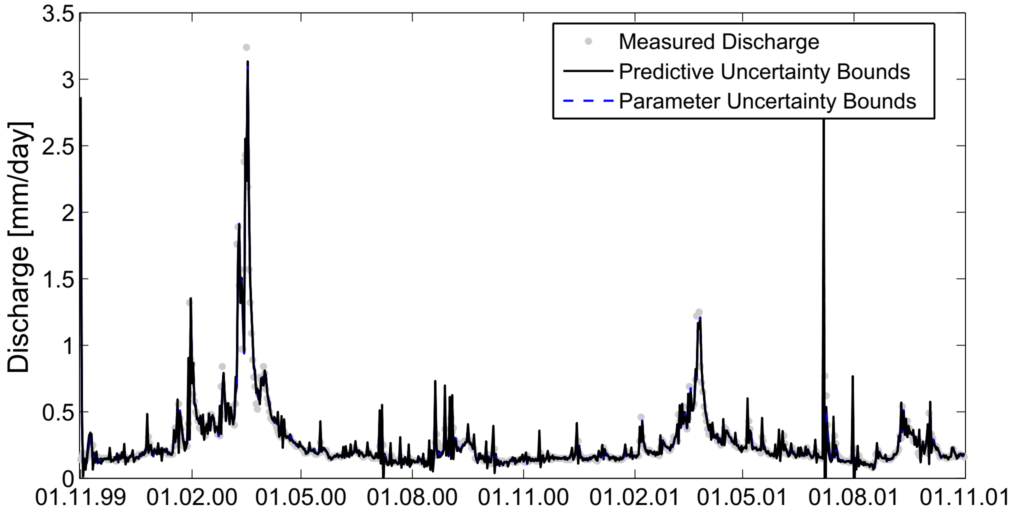


Figure 5.20: The predictive and parameter uncertainty bounds for the Döllnitz catchment using MC^2 with autocorrelation and a 300 m grid resolution.

Table 5.17: The water balance and the Nash Sutcliffe criteria for the Döllnitz catchment using the MC^2 method with autocorrelation for the Validation Period.

	Water Balance (192.6 mm)	Nash Sutcliffe
1000	197.1	0.47
500	194.9	-0.3
300	198.6	0.37
100	198.1	0.19

5.2.3 Meso Scale Gera-Langenberg Catchment

5.2.3.1 The GLUE Methodology

In Table 5.18 are the parameter values, the water balance and the Nash Sutcliffe criteria listed. Almost all parameter lie in the same dimension as the calibrated parameters, with only one exception of T_{kor} for the 6000 m grid aggregation. The dotty plot for the 6000 m grid resolution, Figure 5.22, shows that for the T_{kor} parameter a good Nash Sutcliffe criteria is obtained over the whole feasible range of the parameter. Noticeable is the sudden decrease in the Nash Sutcliffe values (though still within the acceptable range) for the parameter m between the values 0.1 and 0.15, indicating a large sensitivity to this parameter. On the other hand the decrease in the parameter SH_{max} is smooth over the whole range. The largest values of the parameter lie near the upper bound of the possible range, suggesting that other feasible parameter values could be located outside this range. For the 2000 m grid aggregation, Figure A.4 page 123, a similar picture is seen for both

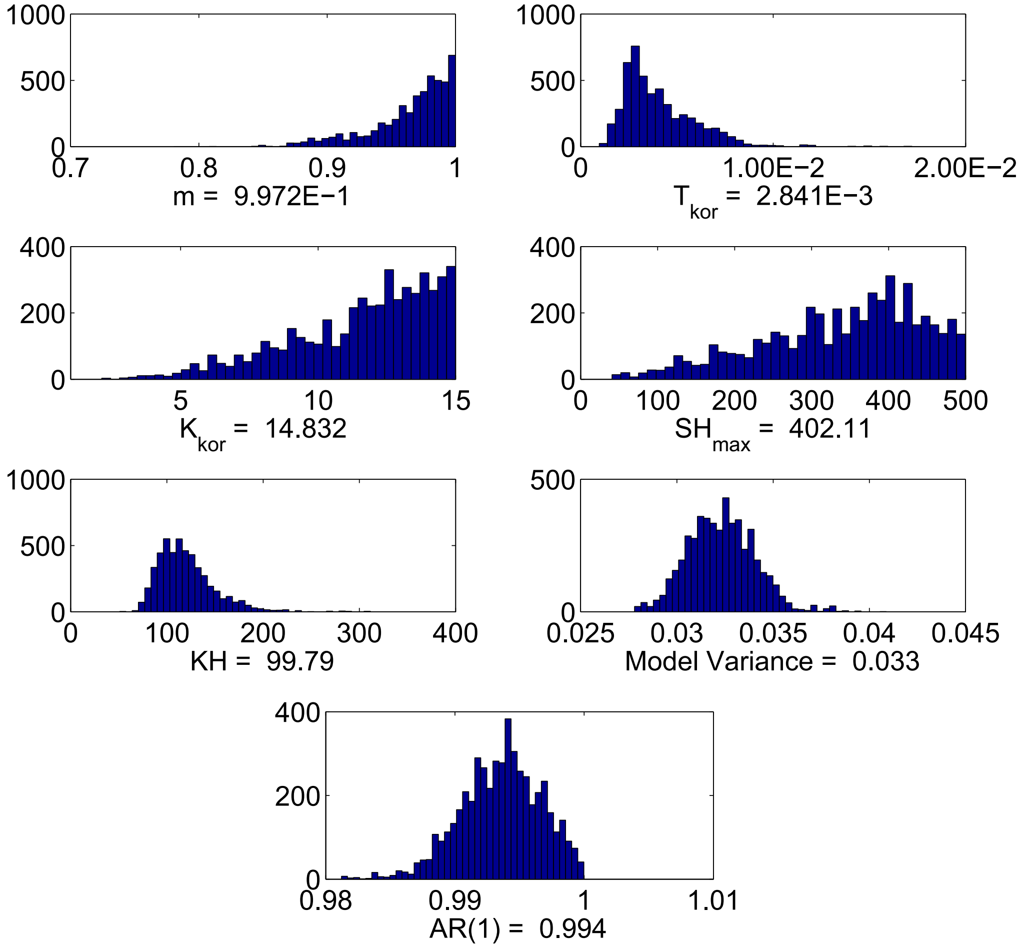


Figure 5.21: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method with autocorrelation and the 300 m grid resolution.

Table 5.18: The water balance, the Nash Sutcliffe criteria and the parameter values for the Gera-Langeberg catchment using GLUE.

	6000 m	4000 m	2000 m	1000 m	500 m
Water Balance (430 mm)	319.8	310.0	314.4	324.1	283.3
Nash Sutcliffe	0.66	0.70	0.73	0.72	0.64
m	0.090	0.128	0.083	0.087	0.085
T_{kor}	8.50E-1	3.03E-3	3.48E-3	3.78E-3	2.26E-3
K_{kor}	8.25	6.85	1.92	9.66	2.80
SH_{max}	482.72	478.58	436.25	489.39	465.42
KH	309.85	342.41	321.19	206.95	294.08

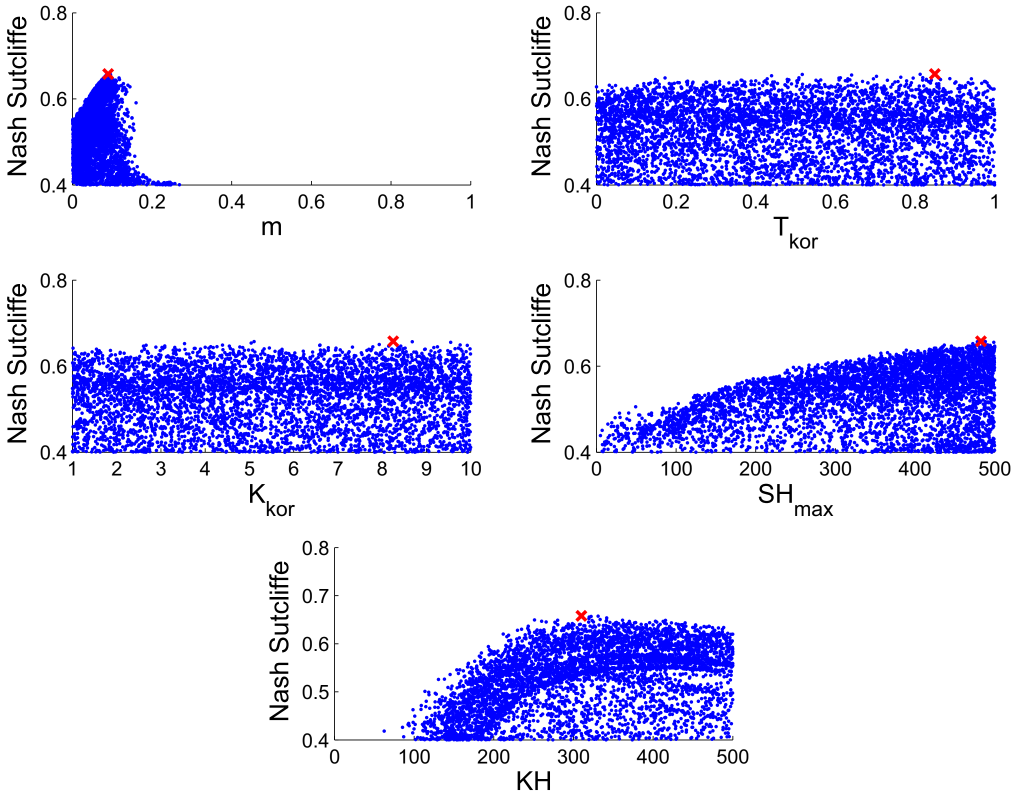


Figure 5.22: The dotty plots for the Gera-Langenberg catchment using a 6000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe

parameters, m and SH_{max}. A different result is noticed for the T_{kor}, a sudden increase in the Nash Sutcliffe values at the lower limit of the range is seen. This is also true for the rest of the resolutions, see Figures A.3 till A.6 in page 123 to 124. The reason for this is the coarse grid aggregation at the 6000 m.

Table 5.18 shows that the water balance is underestimated, with an error ranging from 25% till 34%, this is despite the "relatively" high Nash Sutcliffe between 0.64 and 0.73. To explain this, a look at the hydrograph is needed. Figure 5.23 shows that the baseflow discharge is not captured at all, and the dynamic at low to moderate discharge rate is not represented by the WaSiM-ETH model. The dynamic in the winter periods is sufficiently reproduced, and the peaks are good estimated. This explains the high error in the water balance but the good Nash Sutcliffe.

The results seen in the hydrograph is reflected in the uncertainty quantile. This is clearly seen in Figures 5.24 and A.17 in page 129 for the 6000 m and 2000 m respectively. The uncertainty quantiles for the baseflow do not contain the measured discharge. The reason for this, as mentioned before, is the fact that

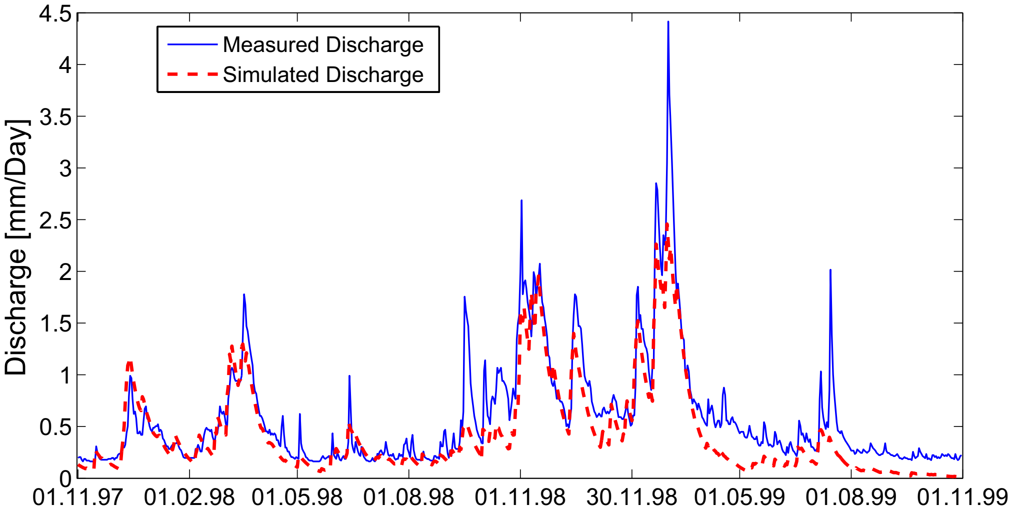


Figure 5.23: The measured and simulated discharge for the Gera-Langenberg catchment using GLUE and a 2000 m grid resolution.

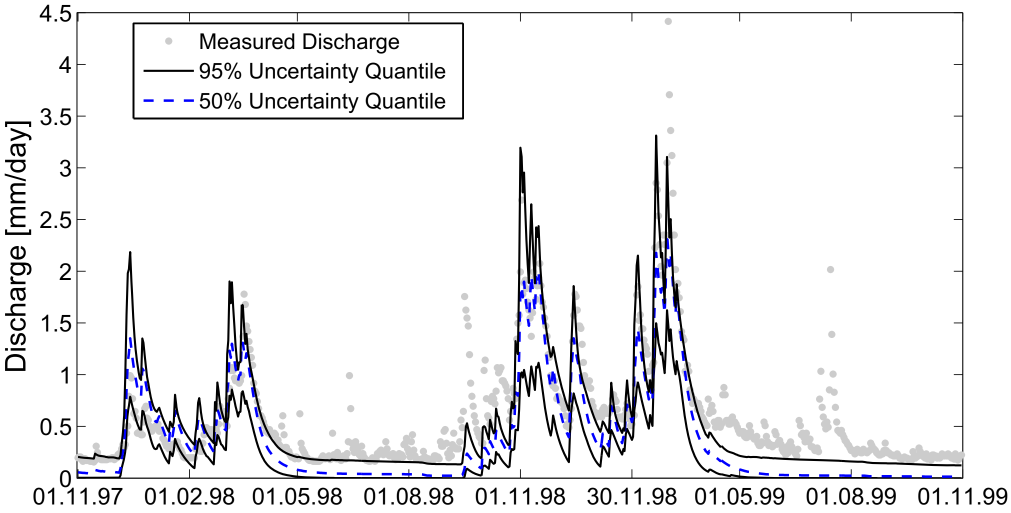


Figure 5.24: GLUE uncertainty bounds for the Gera-Langenberg catchment with resolution of 6000 m.

the WaSiM-ETH model is not capable of representing the dynamic in the system. The same is true for the three summer events. Taking into consideration only the winter months, a small difference is noticed between the 6000 m and the 2000 m grid resolution. The uncertainty quantile for the 2000 m are narrower than for the 6000 m. Implying that the model obtains, at least for the winter events, with the 2000 m better results. Comparing all grid resolutions, it is observed that the for the grids aggregations 500 m (Figure A.19), 1000 m (Figure A.18) and 2000 m comparable results are obtained regarding the uncertainty quantiles, also for the 4000 m (Figure A.16) and the 6000 m. A jump in the uncertainty quantiles is noticed. This could be explained when looking at the entropy in the system. Table 5.19 shows the entropy in the system for the digital elevation model and the topographical index. A significant reduction in the information content of the grids compared to the original grids is seen.
Though the GLUE method in the calibration period provide good results, a different picture is seen in the validation period, negative Nash Sutcliffe values and a mean error of 25% in the water balance. This is seen in Figure 5.20. The dynamic in the system is good captured, nevertheless none of the peaks are met.

Table 5.19: Information content and information content loss for the different grid resolution for the Gera-Langenberg catchment.

	Digital Elevation Model		Topographical Index	
	Information Content	Error	Information Content	Error
25	2.6309		0.7189	
500	2.6115	0.47	0.7048	1.96
1000	2.5430	3.34	0.6506	9.50
2000	2.3182	11.88	0.5596	22.16
4000	2.0647	21.52	0.4255	40.81
6000	1.7268	34.36	0.3928	45.36

Table 5.20: The water balance and the Nash Sutcliffe criteria for Gera-Langenberg using the GLUE method for the validation period.

	Water Balance (398.8 mm)	Nash Sutcliffe
6000	450.6	-0.45
4000	459.1	-0.155
2000	463.7	-0.234
1000	472.7	-0.95
500	430.8	-0.1

This explains the negative Nash Sutcliffe, Figures A.15 till A.11 from page 128 to 127. No differences between the various grid resolutions are found.

5.2.3.2 The MC^2 Methodology

MC^2 without Autocorrelation:

Beside the five chosen parameter an extra statistical parameter (model variance) was introduced in the uncertainty analysis. In Table 5.21 are the parameter values for each grid aggregation. It is noticed that the parameter m values lie within the same order of magnitude, with relatively the same value (ranging from 0.11 to 0.18). The value of T_{kor} decreases as the resolution of the grid becomes finer, with an order of magnitude change of 10^3. This indicates a dependence of this parameter on the grid aggregation for this catchment. In equation 4.5 the parameter T_{kor} is used to estimate the saturation deficit of each cell in the catchment, which plays a role in estimating the surface runoff. The parameter T_{kor} is associated with α_t, specific catchment area per unit length of a grid cell; this is the area draining through one meter of the edge of a grid cell, with β_t the slope angel and with the γ parameter, the mean topographic index. Table 5.22 shows the mean topographic index for the different grid resolutions. As seen, the

Table 5.21: The water balance, the Nash Sutcliffe criteria and the parameter values for the Gera-Langenberg catchment using MC^2 without autocorrelation.

	6000 m	4000 m	2000 m	1000 m	500 m
Water Balance (430 mm)	232.96	232.11	371.63	330.45	332.59
Nash Sutcliffe	0.168	0.311	0.8373	0.752	0.776
m	0.188	0.174	0.119	0.125	0.114
T_{kor}	0.356	0.142	1.75E-02	2.77E-03	7.40E-04
K_{kor}	5.86	2.63	7.29	2.62	7.30
SH_{max}	73.77	496.83	352.68	371.31	376.21
KH	227.0	336.2	251.2	286.6	277.9
Model Variance	0.252	0.202	0.069	0.073	0.073

Table 5.22: The mean topographical index for the Gera-Langenberg catchment for each grid resolution.

	Mean Topographical Index
500 m	17.22
1000 m	18.91
2000 m	20.53
4000 m	22.00
6000 m	22.81

mean topographic index increase with finer grid resolution. Since the parameter m has almost the same value for all grid aggregations, the log term in equation 4.5 should be the same. Since the mean topographic index is increasing, the transmissitivity correction factor T_{kor} should become small to balance the term. The posterior parameter distribution for the 2000 m grid resolution, in Figure 5.25, shows that all parameters have converged. Furthermore, the values of the

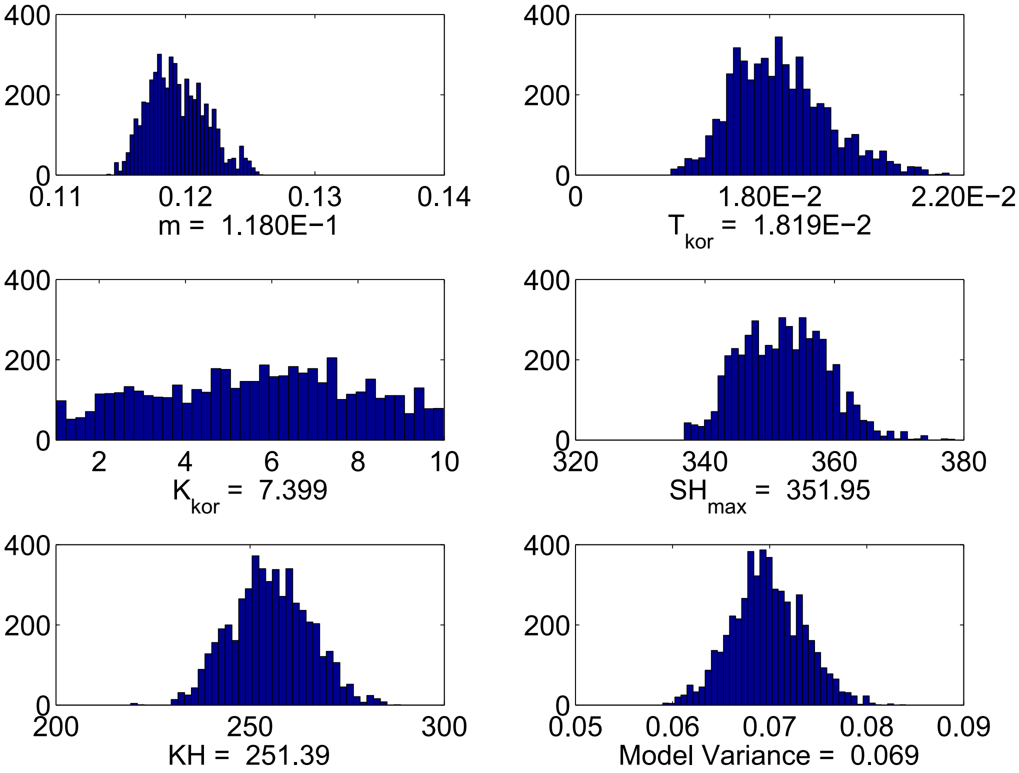


Figure 5.25: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method without autocorrelation and the 2000 m grid resolution

statistical parameter q_{val} decrease with finer grid resolution, implying a better prediction power of the finer grid resolutions (up to 2000 m). This is observed when studying the uncertainty prediction bound seen in Figure 5.26 for the 2000 m and in Figure A.24 page 133 for the 6000 m grid resolutions. The uncertainty bounds for the 2000 m is more close to each other and they match the dynamic in the system. On the contrary to that, the 6000 m bounds are much wider and they do not capture the dynamic in the system, especially for the low flow periods. This is clearly seen, when looking at the discharge curve for both grid resolution in Figures 5.27 and A.28. Moreover, the parameter uncertainty bounds are for both cases narrow. This is due to the low value of the model variance parameter.

In the validation period, on the contrary of what is to expect, the MC^2 has

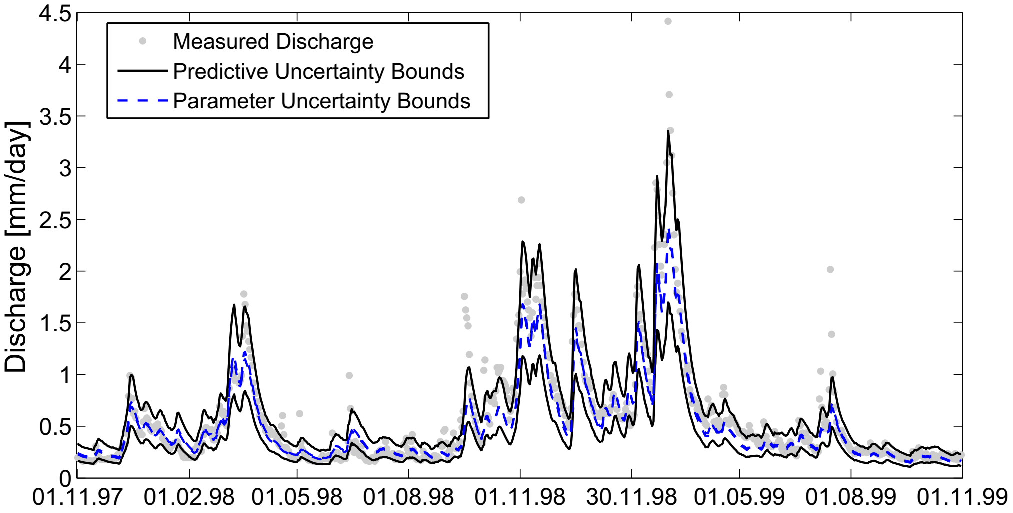


Figure 5.26: The predictive and parameter uncertainty bounds for the Gera-Langenberg catchment using MC^2 without autocorrelation and a 2000 m grid resolution.

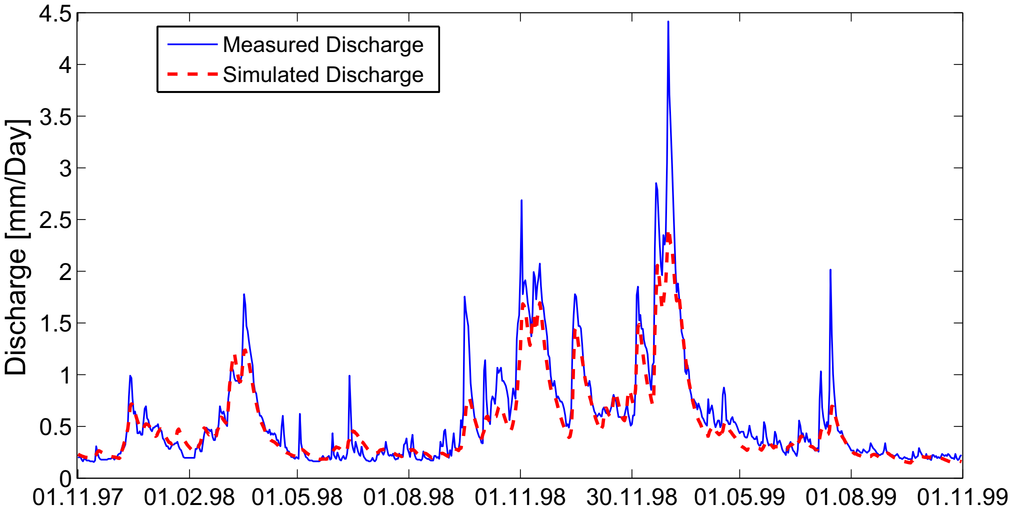


Figure 5.27: The measured and simulated discharge for Gera-Langenberg catchment with MC^2 and a 2000 m Grid Resolution.

achieved good results. These are seen in Table 5.23. The water balance is overestimated with a minimum of 10%, the Nash Sutcliffe values are positive, and only two values (grid resolutions 6000 m and 4000 m) are above the threshold (>0.4). Noticeable is the fact that in contrast to the two other catchments, the MC^2 achieved better results than the GLUE method. This confirms the fact that, the combination of the parameter is the crucial criteria and that there exists no one single optimal parameter set. Figures A.36 till A.32 in pages 137 to 136, show the hydrographs for the catchment. The dynamic in the system is good reproduced. The higher Nash Sutcliffe values for the two grid coarse could be explained by examining the hydrographs. In comparison to the other grid resolutions, more peaks are met, thus improving the Nash Sutcliffe value.

Table 5.23: The water balance and the Nash Sutcliffe criteria for Gera-Langenberg catchment using the MC^2 without autocorrelation for the validation period.

	Water Balance (398.8 mm)	Nash Sutcliffe
6000	438.9	0.45
4000	436.6	0.61
2000	491.5	0.21
1000	501.9	0.21
500	502.4	0.2

MC^2 with Autocorrelation:

Another statistical parameter is added to the Monte Carlo Markov Chain method, which is the autocorrelation parameter. Table 5.24 shows the parameter values. For the autocorrelation parameter the values ranges from 0.8 till 0.99, implying

Table 5.24: The water balance, the Nash Sutcliffe criteria and the parameter values for the Gera-Langeberg catchement using MC^2 with autocorrelation.

	6000 m	4000 m	2000 m	1000 m	500 m
Water Balance (430 mm)	420.6	423.2	421.9	420.2	419.1
Nash Sutcliffe	0.92	0.93	0.95	0.95	0.9
m	0.202	0.107	0.120	0.100	0.580
T_{kor}	0.332	0.278	0.031	0.030	0.460
K_{kor}	4.8	6.2	4.4	7.0	6.9
SH_{max}	36.7	26.8	415.4	485.4	210.6
KH	485.9	484.2	203.6	171.7	449.7
Model Variance	3.30E-2	3.25E-2	2.77E-2	2.75E-2	2.74E-2
AR	0.951	0.966	0.809	0.825	0.798

that 80% to 99% of the error in the discharge data is explained using only one time step. The model variance is smaller than the model variance without autocorrelation. Indicating, that the model variance (without autocorrelation) have captured a part of the error in the discharge data. The model parameters vary strongly within the range, e.g the parameter m values ranges between 0.11 and 0.98. The reason for this is clearly seen when looking at the posterior parameter distributions, Figure 5.28 for 2000 m and Figure A.37 for 6000 m. As expected, the addition of the autocorrelation factor did not only improve the results in the calibration period but also in the validation period. In Table 5.25 are the Nash Sutcliffe values and the water balace for the catchment in the validation period. The hydrographs, Figures A.49 to A.45 in pages 144 to 143 show almost a perfect match between the measured and the simulated discharge. One except is the 500 m grid resolution, where a clear overestimation of the hydrograph is seen. The

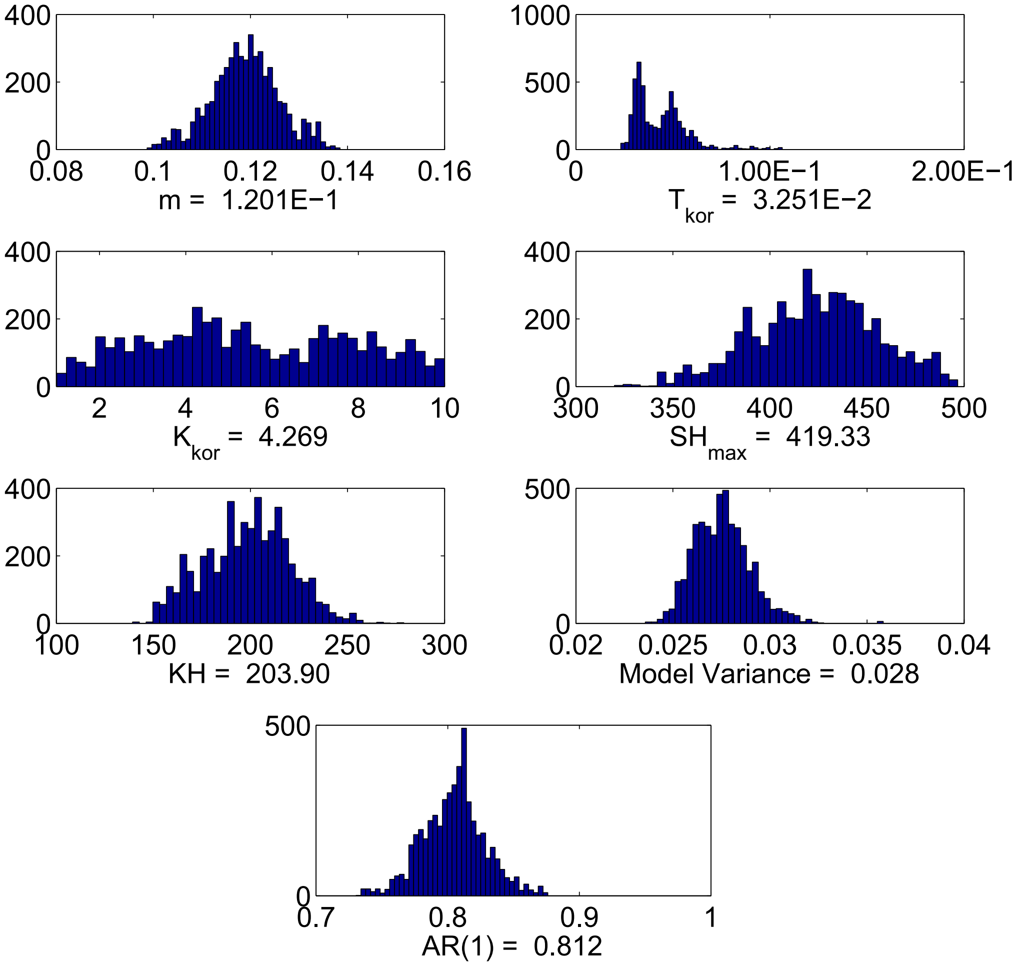


Figure 5.28: Parameter Posterior Distribution based on the MC^2 Methodology with Autocorrelation for the 2000 m Grid Resolution for Gera-Langenberg

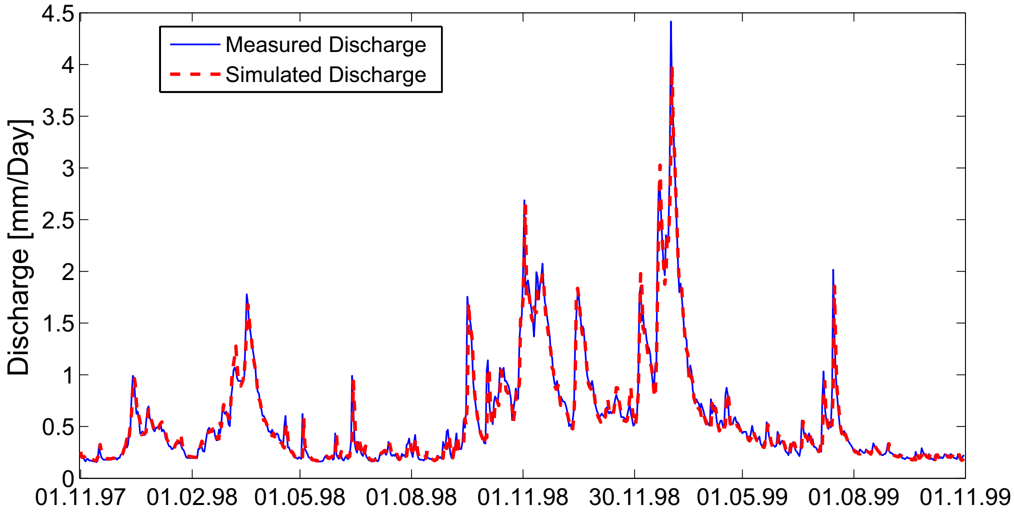


Figure 5.29: The measured and simulated discharge for Gera-Langenberg catchment with MC^2 with autocorrelation and a 2000 m grid resolution.

Table 5.25: The water balance and the Nash Sutcliffe criteria for Gera-Langenberg catchment using the MC^2 with autocorrelation for the validation period.

	Water Balance (398.8 mm)	Nash Sutcliffe
6000	402.1	0.62
4000	398.5	0.72
2000	420.0	0.87
1000	417.1	0.87
500	484.5	0.7

reason for this is the fact that the parameter did not converge at the end of the uncertainty analysis procedure, see Figure A.40 page 140.

5.3 Uncertainty Analysis - Rainfall Uncertainty

5.3.1 Small Scale Lower Mountain - Weida Catchment

For the first two uncertainty analysis methods, the uncertainty in the input data was not taken into consideration. Thus, a new concept was developed and implemented. In Figure 5.30 are the 95% quantiles of the total uncertainty (dark line) and the 95% quantiles due to the input, i.e. rain uncertainty (red line) for the 200 m grid resolution. The total uncertainty quantile has the same meaning as in GLUE methodology, except that they incorporate the rainfall uncertainty due to measurement error. Again the total uncertainty quantiles are narrow and

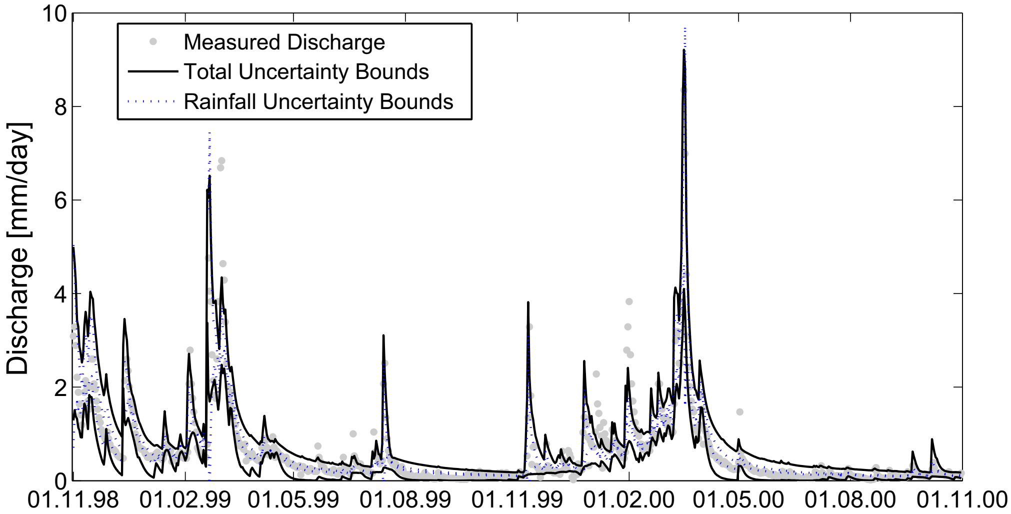


Figure 5.30: The total and the rainfall uncertainty bounds for Weida catchment using INPUT concept and a 200 m grid resolution.

comparable with total uncertainty quantiles from the GLUE methodology. The important part of this Figure is the uncertainty bounds due to rainfall measurement error. These quantile were estimated using Equations 3.31 and 3.36.

Noticeable is the fact that the uncertainty bound due to rainfall measurement errors are wider at major rainfall events and with higher discharge flow. For low discharge values, the rainfall uncertainty does not play an important role. It is also noticed that for the two major peaks, the rainfall uncertainty exceeds the total uncertainty bounds. This shows that the rainfall uncertainty plays an important role in estimating the peak. Since the rainfall uncertainty is large, the model tries to compensate and reduce this uncertainty through the model parameters. The result is that the model underestimates the two peaks. Comparing the other grid aggregations with the 200 m grid resolution hardly any differences appears. In Table 5.26 are the best parameter values from the first run. A relative good water balance for all grid resolutions is achieved, the same for the Nash Sutcliffe criteria. The parameter values lie also in the same order of magnitude. For the 200 m grid aggregation two best parameter sets are found. The parameter values, besides the parameter (m) and (KH), are almost identical. From the GLUE results, it is known that the parameter (m) is the controlling parameter of the model. So with almost double the value of the parameter (m), the same value of Nash Sutcliffe is obtained. This indicates a large influence of the rainfall uncertainty in the system. This could be clearly seen when calculating the Nash Sutcliffe criteria for the second run. In other words, fixing the parameter to the best parameter set, and sampling only from the rainfall. Figure 5.31 shows this results for the 200 m grid aggregation. Though all points lie above the accep-

Table 5.26: The water balance, the Nash Sutcliffe criteria for the Weida catchment using the INPUT concept.

	500 m	300 m	200 m	200 m	100 m
Water Balance (471 mm)	469.59	483.19	491.05	499.56	474.44
Nash Sutcliffe	0.79	0.79	0.79	0.79	0.80
m	4.34E-3	1.66E-3	8.95E-3	1.62E-2	9.59E-3
T_{kor}	2.77E-2	3.39E-2	7.92E-2	8.38E-2	6.53E-2
K_{kor}	6.05	5.83	2.07	2.02	1.81
SH_{max}	44.6	46.5	48.4	40.5	45.7
KH	214.2	232.1	261.5	155.3	252.3

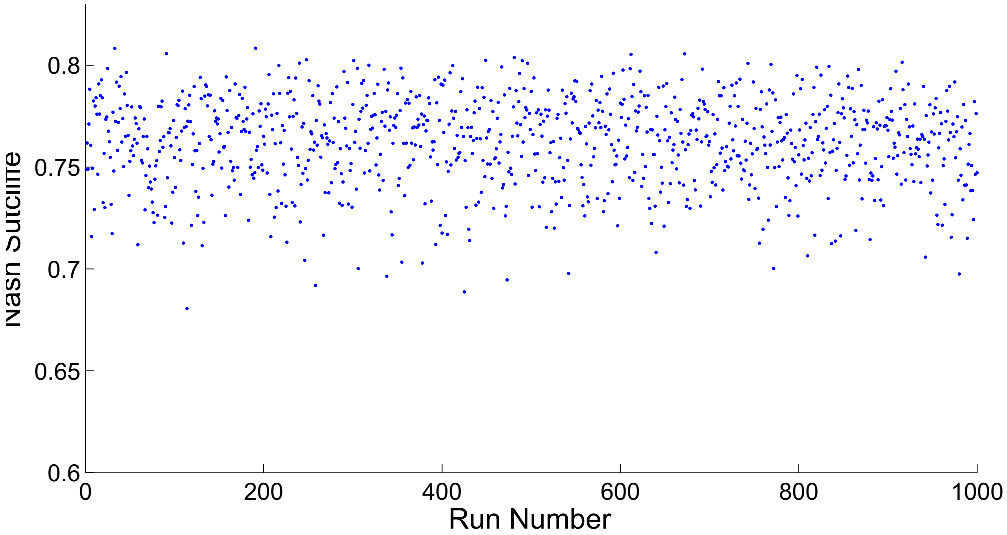


Figure 5.31: The Nash Sutcliffe criteria based on the the uncertainty in the rainfall input data for Weida catchment using the INPUT concept and a 200 m grid resolution.

tance threshold (Nash Sutcliffe >0.6) the variation in the Nash Sutcliffe criteria is relatively large, between 0.67 and 0.82. This variation is caused only by rainfall uncertainty. This implies the important role of the rainfall uncertainty for this catchment. Another consequence is that a non-behavioral model could become behavioral when taking rainfall uncertainty into consideration. The dotty plots, in Figure 5.32, show the same pattern as from GLUE. The simulated discharge hydrograph, seen in Figure 5.33, describes the dynamic of the system in a good manner.

At the validation period the INPUT method preforms as good as the GLUE method, see Table 5.27. The hydrograph, seen in Figures B.58 to B.56 from page 179 to 179, display the dynamic of the system captured by the method,

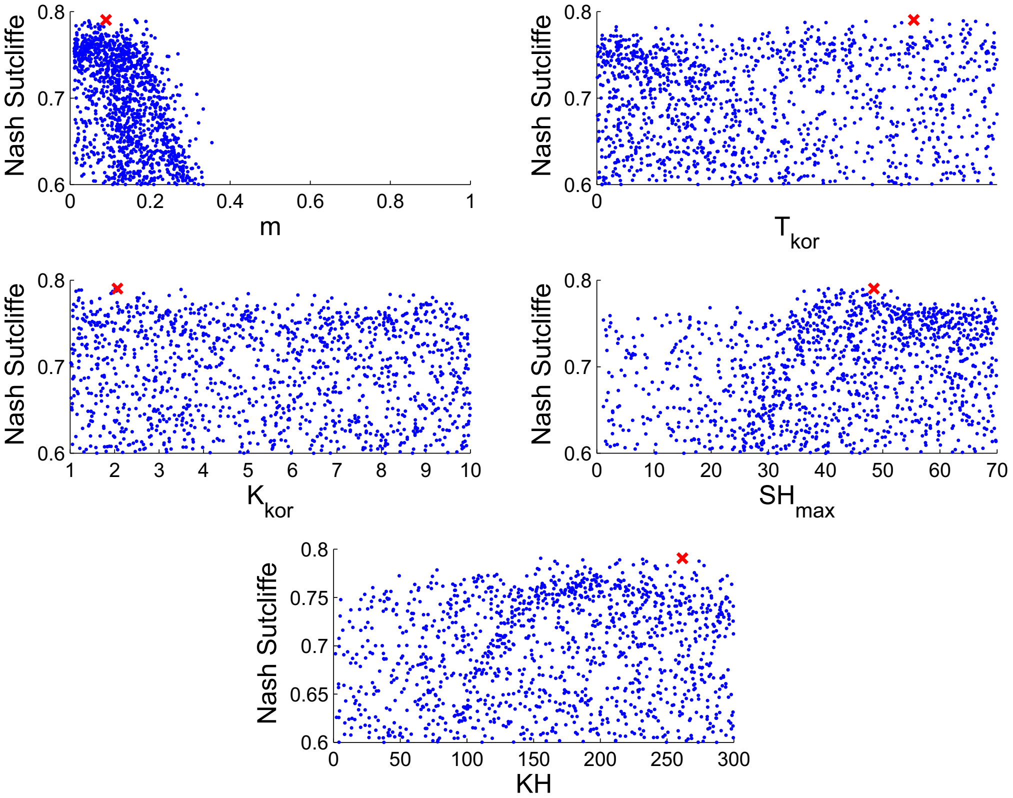


Figure 5.32: The dotty plots for the Weida catchment using the INPUT concept and a 200 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

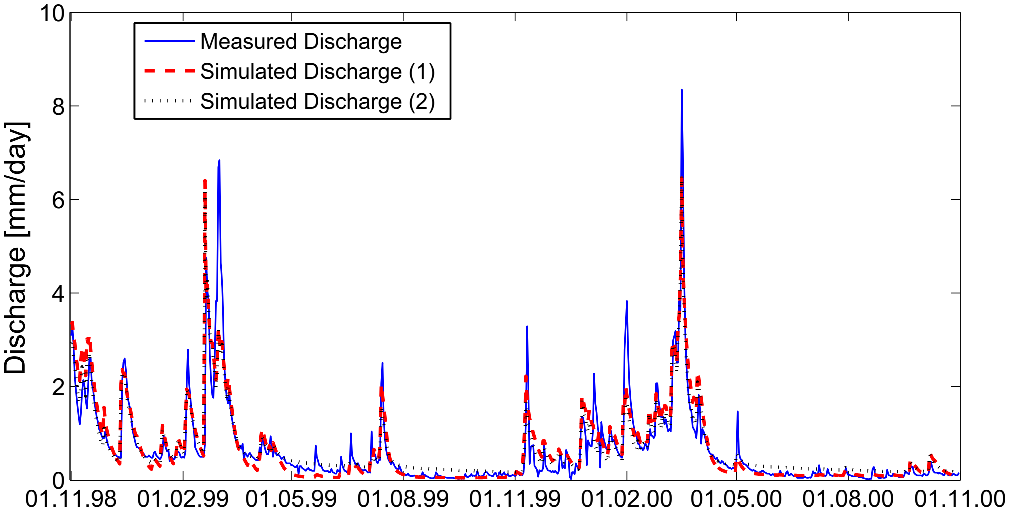


Figure 5.33: The measured and simulated discharge for the Weida catchment with a grid resolution of 200 m based on the INPUT Concept

Table 5.27: The water balance and the Nash Sutcliffe criteria for the Weida catchment using the INPUT method for the validation period.

	Water Balance (351 mm)	Nash Sutcliffe
500	379.1	0.66
300	397.6	0.64
200	393.2	0.66
200	363.4	0.62
100	366.2	0.67

though same as GLUE the major peaks are not well met. Noticeable is the water balance result for the 200 m grid aggregation, though the two best parameter sets did not vary at the calibration period, see Table 5.26, their water balance in the validation period differs. This indicates that one parameter set capture the physical process in both calibration and validation periods, and one does not. The only difference in both parameter sets, as explained earlier, is the value of the parameter (m), emphasizing the importance of this parameter.

5.3.2 Small Scale Loess Catchment - Döllnitz Catchment

To incorporate the rainfall uncertainty in this study for each of the daily rainfall values of the five rain gauges, a normal distribution was generated and samples were taken from this generated distribution, as explained above. Table 5.28 shows the parameter values, Nash Sutcliffe criteria and the water balance of the catchment. Noticeable are the low Nash Sutcliffe values, all below the threshold for behavioral models (0.4). For this reason the threshold was adjusted and all models having a Nash Sutcliffe larger than zero (>0) will be considered as behavioral. Another point, is the value for the parameter SH_{max}. It lies within

Table 5.28: The water balance, Nash Sutcliffe criteria and the parameter values for the Döllnitz catchment using the INPUT concept.

	1000	500	300	100
Water Balance (196.5 mm)	292.58	238.53	275.38	269.99
Nash Sutcliffe	0.16	0.29	0.29	0.26
m	7.36E-2	6.20E-2	3.60E-2	1.58E-2
T_{kor}	0.358	0.906	0.324	0.274
K_{kor}	3.28	1.09	4.88	9.92
SH_{max}	489.1	377.4	253.2	363.5
KH	317.3	428.8	953.6	928.6

the upper half of the feasible parameter range, in contrast to the values found by the GLUE methodology.
Figure 5.34 shows rather a gloomy dotty plots for the grid resolution 300 m.

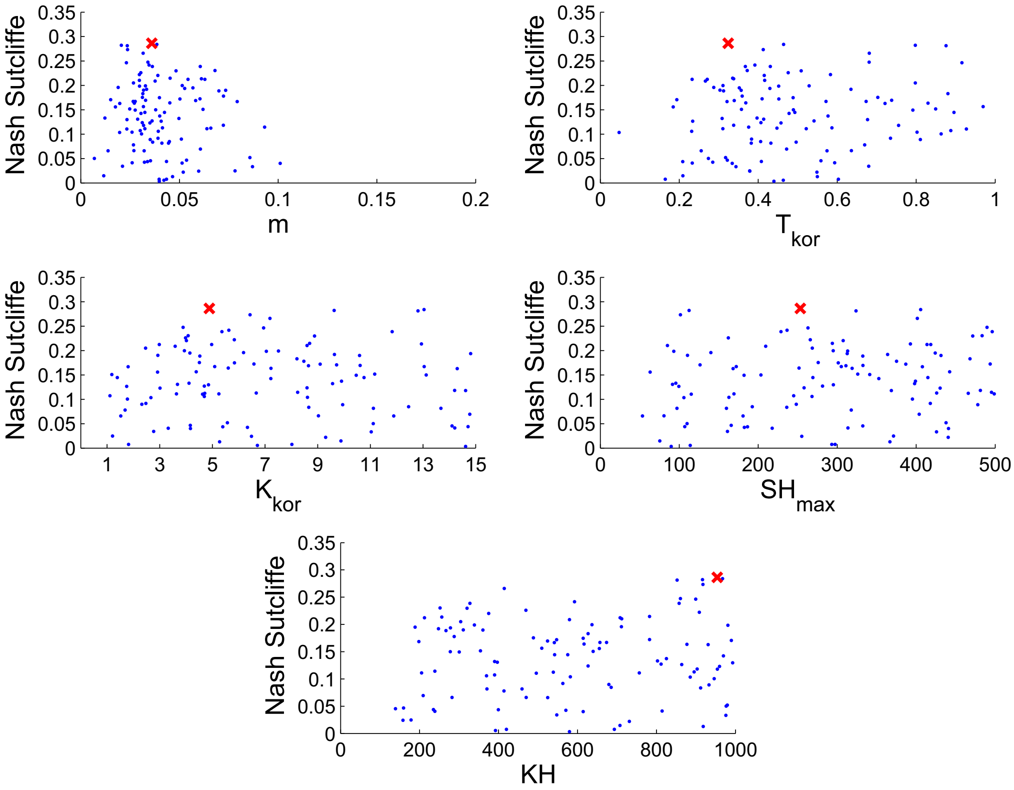


Figure 5.34: The dotty plots for the Döllnitz catchment using a 300 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

There exists a rather small number of behavioral models compared with GLUE, despite the adjustment of the threshold, see Figure 5.13. One reason for this small number of behavioral models is the strong interaction between input and model parameters. In other words, a small change in the rainfall value could lead to an underestimation or an overestimation of the discharge from the catchment, and keeping in mind that the Nash Sutcliffe criteria is highly sensitive to outliers, this combination leads to the fact that most models will not make it as behavioral models.
Despite the small number of behavioral models, the total uncertainty quantile and the rainfall uncertainty quantile are calculated. Figure 5.35 shows the uncertainty quantiles for the 300 m grid aggregation. Again the WaSiM-ETH model is not capable of representing the catchment dynamic in a sufficient and satisfactory way. This is reflected in the uncertainty quantiles. The rainfall uncertainty quantiles are only significant for the two major spring events in the years 2000

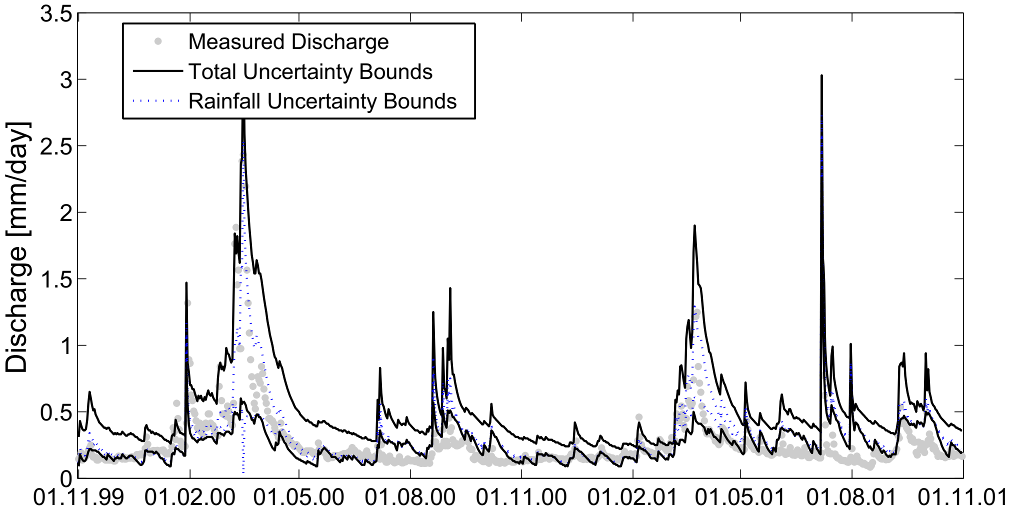


Figure 5.35: The total and the rainfall uncertainty bounds for the Döllnitz catchment using INPUT concept and a 300 m grid resolution.

and 2001. For the rest of the calibration period, the rainfall uncertainty quantiles are negligible. For the summer event 2001 (08.07.01) both uncertainty quantiles are identical, and represent one line. This supports the theory of the WaSiM-ETH model not able to simulate this event, or capture the physical process behind it. Depending on the explanation above, why a small number of behavioral models are obtained, one expects that the rainfall uncertainty is relatively large, but looking at Figure 5.36 it is not the case. One reason for this, the fact that the best behavioral model is the best parameter combination that suppresses or can handle uncertainty in the rainfall data. Meaning that, the parameter combination can smooth the error in the rainfall measurement data. The INPUT method achieved also positive Nash Sutcliffe values, at least for the grid aggregations 300 and 100 m, though the values are still below the acceptance threshold (>04), see Table 5.29. The water balance, on the other hand, exhibit an average error of about 20%. The discharge curves are seen in Figures C.41 till C.43 in pages 199

Table 5.29: The water balance and the Nash Sutcliffe criteria for the Döllnitz catchment using the INPUT concept for the calidation period.

	Water Balance (192.6 mm)	Nash Sutcliffe
1000	292.7	-0.61
500	193.4	0
300	233.7	0.21
100	226.6	0.21

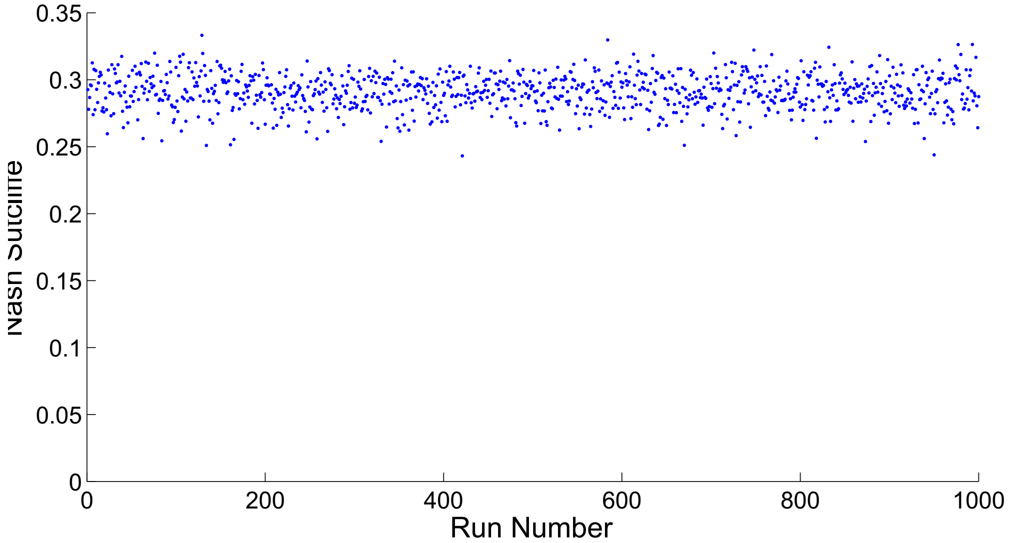


Figure 5.36: The Nash Sutcliffe criteria for the second run using the INPUT Concept for the Döllnitz Catchment using a 300 m grid resolution

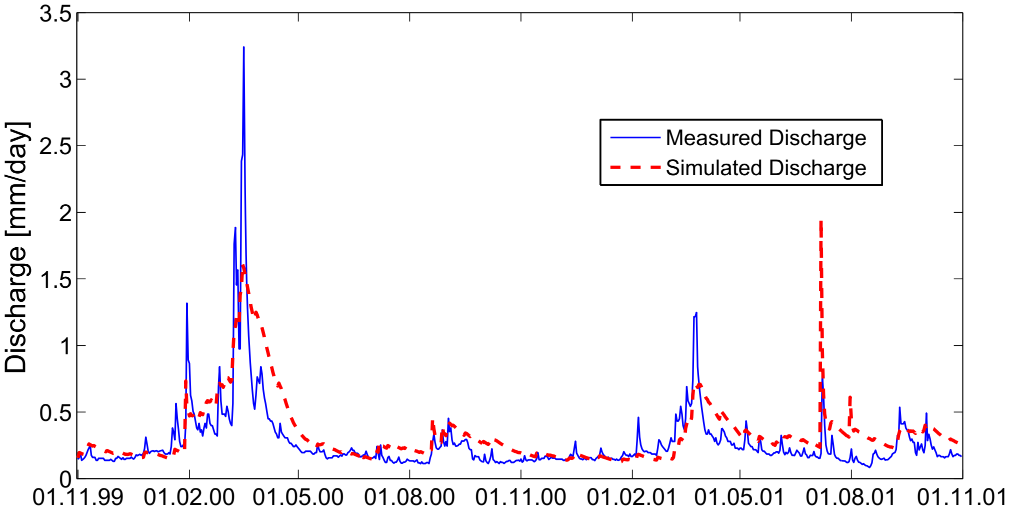


Figure 5.37: The measured and simulated discharge for the Döllnitz catchment using the INPUT concept with a 300 m Grid Resolution.

to 200.

5.3.3 Meso Scale - Gera-Langenberg Catchment

Incorporating the rainfall uncertainty into the uncertainty analysis, a similar picture regarding the uncertainty bounds emerges. Figure 5.38 shows the uncertainty bounds for the 2000 m grid aggregation. The uncertainty bounds due to rainfall measurement error (the red line) are wider for major discharge events (08.12.98 till 13.04.98 and 28.10.98 till 12.12.98) and compromise almost the whole uncertainty bounds. But still, three peak flows are not captured by the model (16.09.98, 03.03.99 and 15.07.99), emphasizing the fact that, the model is not capable of capturing the physical processes behind the formation of these peak flows. This could be considered as model structure uncertainty. This is also true for the rest of the grid resolutions; see Figures A.50 till A.53 from page 145 till 146). Again for some of the grid aggregations it is seen that the rainfall uncertainty for some events are higher than the predictive uncertainty of the model. This is due to the interaction between the model and the input data.

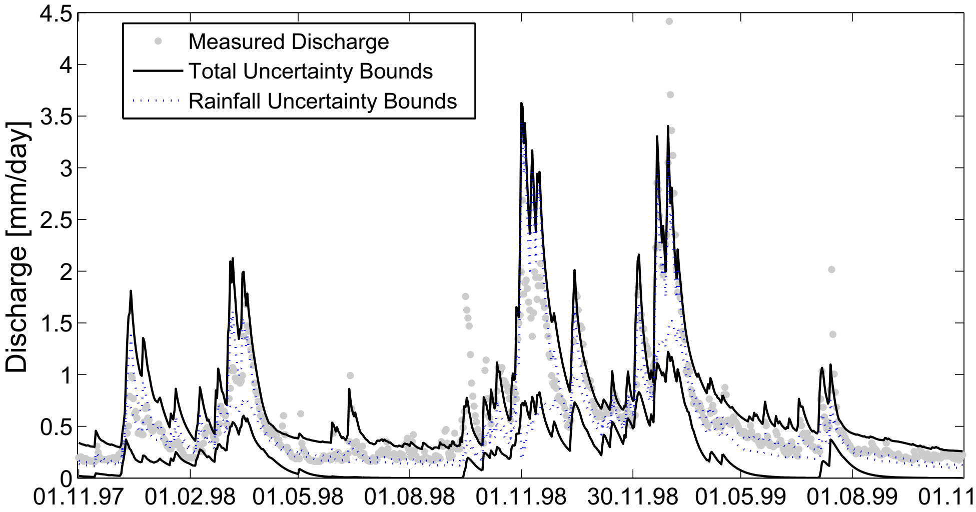


Figure 5.38: The total and the rainfall uncertainty bounds for the Gera-Langenberg catchment using the INPUT concept and a 2000 m grid resolution.

In Table 5.30 are the parameter values, the Nash Sutcliffe and the water balance values for the different grid aggregations. Figure 5.39 shows the dotty plots from the first run. A homogeneous distribution is obtained over the whole parameter range.

A high Nash Sutcliffe was achieved for all grids sizes. The water balance on the other hand is underestimated for all grids, one reason for this is the fact that the model is not capable of reproducing the three peak flows, mentioned above, Figure 5.40. Looking at the results from the second run, a different picture is

Table 5.30: The water balance, the Nash Sutcliffe criteria and the parameter values for the Gera-Langenberg catchment using the INPUT Concept

	6000 m	4000 m	2000 m	1000 m	500 m
Water Balance (430 mm)	339.4	355.5	367.3	357.2	336.9
Nash Sutcliffe	0.79	0.80	0.81	0.79	0.80
m	4.31E-2	6.09E-2	9.33E-2	8.43E-2	4.98E-2
T_{kor}	3.32E-1	3.29E-1	1.26E-1	3.58E-1	1.32E-1
K_{kor}	4.13	6.70	8.09	3.28	9.21
SH_{max}	195.3	452.32	465.65	489.14	318.74
KH	253.35	345.7	253.25	317.35	263.78

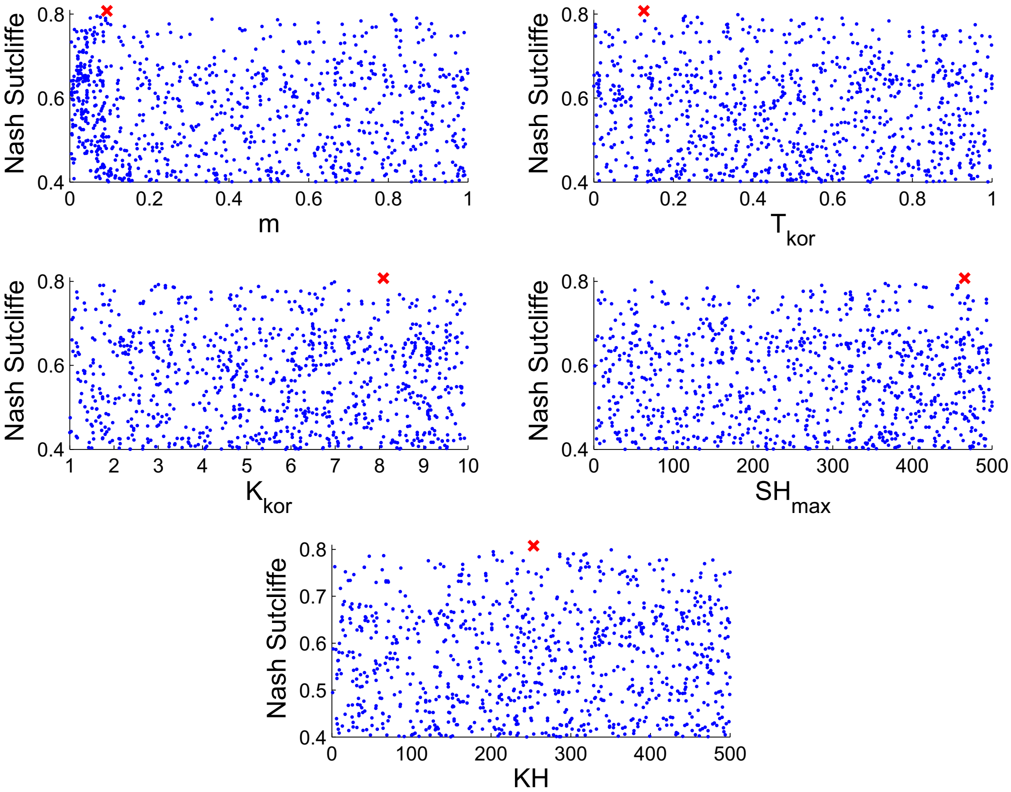


Figure 5.39: The dotty plots for the Gera-Langenberg catchment using a 2000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe

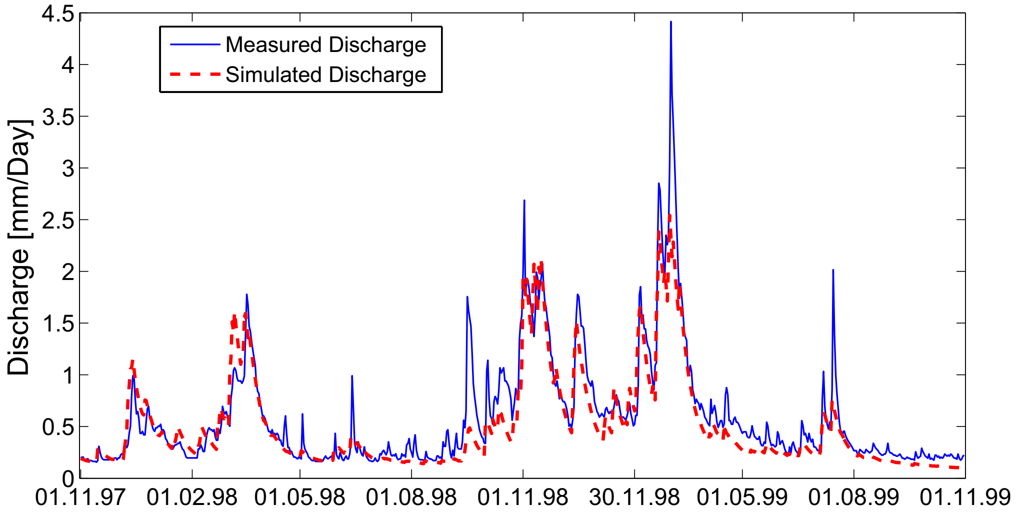


Figure 5.40: The total and the rainfall uncertainty bounds for the Gera-Langenberg catchment using INPUT concept and a 2000 m grid resolution

seen, when compared to the results from the Weida catchment. Figure 5.41 shows the Nash Sutcliffe values of the 1000 runs, when fixing the parameter values and sampling only from the rainfall. These values vary within a small range (between 0.76 and 0.82), indicating a minor influence of the rainfall measurement uncertainty for this large catchment. The INPUT method performs when compared to the other methods in a moderate way in the validation period. Table 5.31 show the Nash Sutcliffe and the water balance of the catchment. In contrast to the

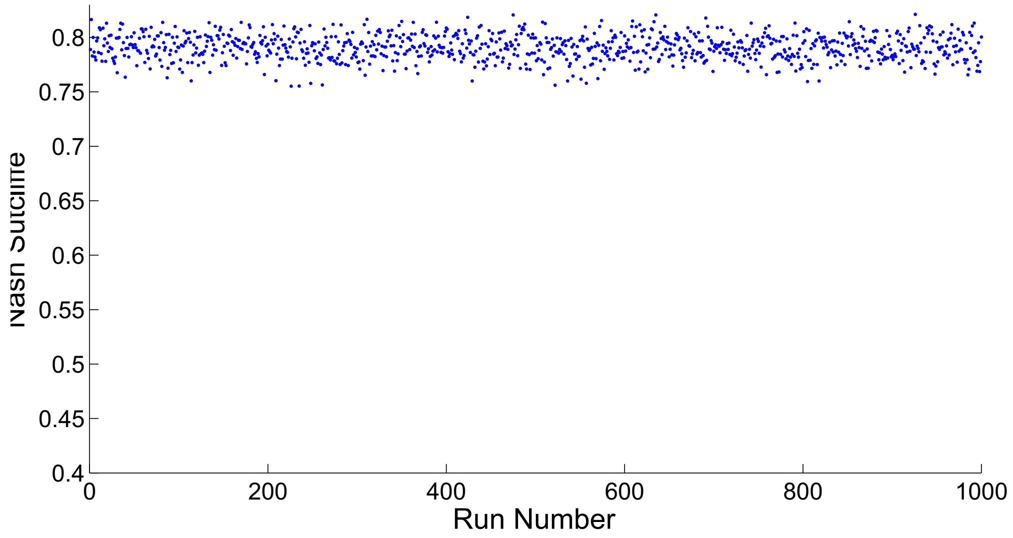


Figure 5.41: The Nash Sutcliffe criteria for the Gera-Langenberg catchment using the INPUT concept and a 2000 m grid resolution.

Table 5.31: The water balance and the Nash Sutcliffe criteria for the Gera-Langenberg catchment using the INPUT concept for the validation period.

	Water Balance (398.8 mm)	Nash Sutcliffe
6000	348.3	0.32
4000	371.6	0.21
2000	367.3	0.25
1000	317.4	0.24
500	349.3	0.23

GLUE method, the INPUT method has achieved positive Nash Sutcliffe values in the validation period, despite the fact that the water balance of the catchment is underestimated. So the consideration of the rainfall measurement error in the overall uncertainty analysis has improved the model predictability.

5.4 Methodology Comparison and General Discussion

After representing the results of the uncertainty analysis using three different methodologies and discussing these results, a closing discussion is made here and a comparison between the methods is undertaken. Basically a direct comparison of the methods is not possible, since the uncertainty analysis is done using different concepts (GLUE and INPUT informal likelihood, MC^2 probability theory / formal likelihood) and the meaning of the uncertainty bounds / quantiles are not comparable. Despite this, a general comparison and the applicability of the methods is made here.

One point, which is common regardless of the method at hand, is the methods do *not* improve the model. In other words, if the model is not able to simulate the catchment and the physical processes behind the generation of the discharge, then the methods can not improve the model. This is clearly seen for the Döllnitz catchment. The WaSiM-ETH model with the TOPMODEL approach is not able to simulate the catchment in a good manner, which was reflected in results of the uncertainty analysis. The TOPMODEL approach was developed for mountainous catchments and the Döllnitz catchment is a flat catchment. This is one disadvantage of the methods at hand, they are not able to give information about the sources of this uncertainty in the model, since they do not differentiate between different sources of uncertainty, except to parameter and predictive uncertainty. One may argue that the MC^2 with autocorrelation has improved the results (both Nash Sutcliffe and the water balance) dramatically, but this was only possible since the error, due to input error and model structure error, in the discharge from the time step before was eliminated using the autocorrelation factor. It did not define or consider a new source of uncertainty in modelling process but it excretes the error of the modelling process found in the simulated discharge from consideration in the next time step. The problem here, as discussed previously, is the method is not applicable if the model is used for predicting of future events, or the evaluation of new scenarios, where no measured discharge is available to calculate the error at any time step.

Though Feyen et al. (2007) in their study did not incorporate the autocorrelation factor in their uncertainty analysis, they discuss that not considering the autocorrelation in the uncertainty analysis questions the reliability of the parameter derived from the posterior parameter distribution for accurate discharge simulation. This again rises from the assumptions made when deriving the likelihood function. For the MC^2 methods, incorporating the autocorrelation give a more reliable parameter estimation, which was also proved, but the applicability of this method to use the best parameter set to extrapolate the discharge simulation is limited. In other words, the method could be only used when discharge data is available. Despite this, the autocorrelation factor AR provides extra information about the system. The autocorrelation factor acts as an indicator, how good the model represent the physical processes in the catchment. A small autocorrelation factor indicates that the error in a certain time step is independent of the error of the error in the time step before. In other words the error at a certain time step is caused mostly from the interaction between the modelling and the input uncertainty. The autocorrelation factor is implemented in order not to overestimate the information content of the singel points. The GLUE and the INPUT method, on the other hand, do not encounter this problem, since they address the error in the system implicitly and do not make any assumptions about it.

When taking the fact that the methods only differentiate between the parameter uncertainty and the predictive uncertainty, the results (Nash Sutcliffe and water balance) obtained from the different methods do not vary at all. The same results were also found by Balin (2004). Though the MC^2 method is capable of gaining good results with a relatively small run numbers. The GLUE methods, on the other hand, gives a better understanding about the interaction between the parameters. With the new method INPUT, the input uncertainty is considered into the overall uncertainty analysis, without forcing or manipulating them to get good simulations. This is one advantage of the new method compared to the other methods, the implicit consideration of the uncertainty in the input data, in the study the rainfall measurment uncertainty. Other methods, which tried to consider uncertainty in the rainfall data, usually introduced a multiplier to manipulate the rainfall data. This multiplier was considered as a parameter in the uncertainty analysis. The INPUT concept implements the uncertainty in the input data (e.g. rainfall data) in the total uncertainty analysis, without changing or manipulating the data. Another advantage is the radical reduction of the computational cost and time. This is done by making some assumptions about the form of the error in the input data. This should be done after thorough thinking, since this assumption will have a large influence on the results of the uncertainty analysis. Despite this, the consideration of the input uncertainty in the total uncertainty analysis improved the predicting power of the model. The parameter uncertainty bounds envelops more measured discharge points, compared to GLUE.

McMillan and Clark (2009) showed in their study that using an informal likelihood function (extended Nash Sutcliffe criteria) has improved the calibration process and the estimation of the peak discharge. Furthermore, they argue the difficulty of using only one statical goodness-of-fit measurment or objective function to summarize the hydrological performance/response. They plead the use of more than one statistical measurement or objective functions. The results obtained in this work imply the same conclusion. Though high Nash Sutcliffe criteria was achieved, the water balance in the catchment was not always good. The incorporation of the water balance in the catchment as a goodness-of-fit criteria should have solved this problem to obtain better results.

After the finding from McMillan and Clark (2009) an informal likelihood function, even if implemented in a Monte Carlo Markov chain, would give more information about the system, since a greater volume of the parameter space is being sampled. Thus providing more information about the response surface, for example the location of the local optima and/or the flat areas. Stedinger et al. (2008) argue that the use of an informal likelihood function (e.g. Nash Sutcliffe criteria) within the GLUE methodology does not distinguish a statistically valid parameter set from a bad fit. They show that using a formal likelihood within GLUE, leads to a better recognition of the statistically valid parameter sets, and also a statistically found uncertainty results. It should be made clear here for the reader, that the issue with formal and informal likelihoods and MC^2 and GLUE is a highly controversial subject in the community. Regardless of this, and based on the results acquired in this work, one could say that both methods have achieved similarly good or bad results (measured using the Nash Sutcliffe and the water balance).

All methods did not find any differences between the different grid resolutions for the Weida catchment. For the Döllnitz and the Weiße Elster catchment, on the other hand, a slight difference for the 1000 m and the 4000 m and 6000 m grid resolutions was found respectively. Brasington and Richards (1998) have found a clear cut, a grid resolution threshold at a resolution between 100 and 200 m. Their results coincide with the degradation in the information content (for the 100 m grid resolution a 4.28% information content loss in comparison to the 20 m and for the 200 m 25% loss) of the digital elevation model. In this study the information content in both the digital elevation model and the topographical index were calculated, as seen in Tables 5.5 page 69 for Weida, 5.12 page 78

for Döllnitz and 5.19 page 90 for Weiße Elster. For the Weida catchment the information content loss due to aggregation amount to only 4.1 % for the digital elevation model and almost 22 % for the topographical index, which explains the fact that no differences between the grid resolutions was found. This corresponds to the finding from Brasington and Richards (1998). For the Döllnitz catchment, a jump in the information content loss is noticed between the 500 m and the 1000 m grid resolution from 2.51 % to 7.4 % for the digital elevation model and from 16.9 % to 33.2 % for the topographical index. Despite this large degradation, no obvious cut or threshold is identified. The reason for this is the fact that the WaSiM-ETH model is not capable to simulate the catchment, as discussed before, in a good manner. On the other hand, the Weisse Elster catchment shows a clear cut between the 2000 m and the 4000 m grid resolution, which consorts with the jump in the information content loss, see Table 5.19 in page 90. The percentage in information content loss is the indicator or the criteria, with which the optimal grid resolution could be found. As a guidance value serves the value 20 % information content loss for the choice of the optimal resolution. Another point is, calibrating the model will compensate the information content loss to a certain degree. This was also found by Brasington and Richards (1998).
Another scale effect was noticed in the parameter values, specially the T_{kor} parameter value. This is due to the aggregation of the digital elevation model and the resulting topographical index. The T_{kor} parameter tries to compensate the information loss. This was also found by Franchini et al. (1996) "The hydraulic conductivity inevitably tends to increase as the grid size increases". This behavior is clearly seen for the Gera-Langenberg catchment, Tables 5.18 and 5.21 in pages 87 and 91. They argue that for large grid resolutions the meaning the hydrological conductivity is artificial, and "incorporates both the actual hydraulic characteristic of the soil and the spatial expansion effect connected with the size of the grid". This is also has been confirmed by Brasington and Richards (1998); QUINN et al. (1995). This is another point in uncertainty analysis or in calibration, the meaning of the calibrated parameter values. The hydrological conductivity is measured at lab scale, and applied to a grid cell in the size of 100*100 m^2. Thus losing any physical meaning.

Chapter 6

Conclusion

The work at hand has three major goal related to recent research topics in the hydrological modelling, which uncertainty analysis method is suitable for my application? how does the interaction between input (different grid resolution / rainfall measurement uncertainty) and parameter looks like? and which role does the catchment characteristics and size plays for the interaction?

Which Uncertainty Method is suitable for my Application?

As explained in section 5.4 both GLUE and MC^2 methods are quite similar to each other despite the different concepts on which they are based. The new method INPUT, provides a promising development of the uncertainty analysis tools. It studies the interaction between the input (e.g. rainfall measurement error) and the model parameters. More work needs to be done to explore the full potential of the method. For example, in this work it was applied to input rainfall measurement error, the method allows to study the effect of rainfall field distribution in the catchment.

When applying any of the uncertainty analysis methods caution should be taken. The MC^2 method assumes a normal, constant variance and time independence simulation error. These assumptions need to be guaranteed, through Box-Cox transformation and autocorrelation. For the new method INPUT, the assumption made about the form of the error in the input data has a large influence on the results obtained from the uncertainty analysis. The introduction of the first order autocorrelation factor improved the results for all catchments considerably, but using this method no forecast can be made.

How does the Interaction between Input (different Grid Resolution / Rainfall Measurement Uncertainty) and Parameter looks like?

For the Weida catchment no differences between the four grid resolutions was established regardless of the method used. The reason for this is the information content loss due to aggregation is relatively low, so that no effect is observed. On the other hand for the Gera-Langenberg catchment, a threshold is observed between the 2000 m and the 4000 m grid aggregation. For the Döllnitz catchment, despite the fact that the WaSiM-ETH is not capable of simulating the catchment, no significant differences was observed between the grid aggregations.

For finding the optimal grid resolution, many factors play an important role, and a balance between them should be found. The grid should be detailed to represent the catchment in a sufficient way, and it should be gross aggregated in order reduce the computational cost. As a guideline serves the percentage of information content loss due to aggregation for determining the optimal grid resolution. Up to 20% information content loss compared to the original resolution could be tolerated without any significant degradation of the simulation results.

Using the INPUT method the interaction between the rainfall measurement uncertainty and the model parameter uncertainty has been observed. The INPUT method allows the integration of the input uncertainty in the overall uncertainty analysis without forcing or manipulating the input data to obtain a good result. Though assumptions need to be made about the form and magnitude of the error in the input data. These (strong) assumptions influence the outcome of the uncertainty analysis, so caution should be taken, when making such assumptions.

It was shown that the model through calibration compensates and counteracts the error in the rainfall data measurement. This finding also confirms the common knowledge that calibration compensates for the other sources of error in the modelling process. In other words, using only the parameter different sources of error and uncertainty (input, model structure, discharge, etc...) should be balanced. Thus emphasizing the easiness to use only the parameter for the calibration and uncertainty analysis and at the same time the difficulty of setting apart the different sources of uncertainty. Another find is the fact, without integrating other sources of information in the modelling process a rather limited and an incomplete pictures is obtained about the physical processes in the catchment.

Which Role does the Catchment Characteristics and Size play for the Interaction?

Both size and characteristics of the catchment play an important role in the uncertainty analysis. The uncertainty analysis results are as good as the model simulation of the catchment. This is clearly seen for the catchment Döllnitz. The WaSiM-ETH model with the TOPMODEL approach was not capable of simulating the catchment, due to its characteristics. Since the uncertainty analysis are done only on the basis of model parameter (GLUE and MC^2), its results are as good as the model results, which are not satisfactory in any means. This could be considered as a downside of such uncertainty analysis methods, or this raises the difficulty of distinguishing and separating the diverse sources of uncertainty. Other effects of the catchment characteristics is seen when considering the autocorrelation factor AR(1). The Weida catchment is characterized with fast response to rainfall events, thus the value of the autocorrelation factor is lower than 1. On the other hand, for the Gera-Langenberg catchment this values is around 1.

For the rainfall measurement uncertainty a different behavior is observed. For the small catchment Weida, the rainfall uncertainty plays an important role, on the contrary to the Gera-Langenberg catchment, large catchment, where the rainfall uncertainty does not have any effects at all. For the large catchment the random errors in the rainfall measurements compensate each others effect. Thus in the sum marginal overall effect is achieved.

One last point regarding uncertainty, uncertainty results should prepared and communicated to the stakeholders. The work of uncertainty analysis does not end in the research, but it should be treated and made understandable for the authorities, Stakeholders and the public.

Bibliography

Ajami, N. K., Q. Y. Duan and S. Sorooshian, 2007: An integrated hydrologic bayesian multimodel combination framework: Confronting input, parameter, and model structural uncertainty in hydrologic prediction. Water Resources Research, **43**. [cited at p. 37]

Anderson, E. A., 1973: National weather service river forecast system - snow accumulation and ablation model. Tech. Rep., National Oceanographic and Atmospheric Administration (NOAA). [cited at p. 49]

Balin, D., 2004: Hydrological behaviour through experimental and modelling approaches. pplication to the haute-mentue catchment. Ph.D Thesis, Ecole Polytechnique Federale de Lausanne. [cited at p. 25, 26, 29, 30, 33, 34, 35, 36, 108, 204]

Beck, M. B., 1987: Water-quality modeling - a review of the analysis of uncertainty. Water Resources Research, **23**, 1393–1442. [cited at p. 2]

Becker, A. and P. Braun, 1999: Disaggregation, aggregation and spatial scaling in hydrological. Journal of Hydrology, **217**, 239–252. [cited at p. 6]

Beven, K., 2001: On stochastic models and the single realization. Hydrological Processes, **15**, 895–896. [cited at p. 12]

Beven, K., 2008: Environmental Modelling: An Uncertain Future? Taylor and Francis. [cited at p. 2, 15, 17, 27, 28, 30]

Beven, K. and A. Binley, 1992: The future of distributed models - model calibration and uncertainty prediction. Hydrological Processes, **6**, 279–298. [cited at p. 2, 13, 25, 27, 28, 204]

Beven, K. and M. J. Kirkby, 1979: A physically based variable contributing area model of basin hydrology. Hydrol. Sci. Bull., **24**. [cited at p. 47, 53]

Brasington, J. and K. Richards, 1998: Interactions between model predictions, parameters and dtm scales for topmodel. Computers & Geosciences, **24**, 299–314. [cited at p. 6, 7, 60, 109, 110]

Braun, L. N., 1985: Simulation of snowmelt-runoff in lowland and lower alpine regions of switzerland. Tech. Rep., ETH Zurich. [cited at p. 49]

Bruneau, P., C. Gasuelodoux, P. Robin, P. Merot and K. Beven, 1995: Sensitivity to space and time resolution of a hydrological model using digital elevation data. Hydrological Processes, **9**, 69–81. [cited at p. 6]

Doherty, J., 2004: Pest - model-independent parameter estimation, user manual: 5th edition. Tech. Rep., Watermark Numerical Computing. [cited at p. 55, 57, 58]

Duan, Q. Y., S. Sorooshian and V. Gupta, 1992: Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research, **28**, 1015–1031. [cited at p. 55]

Dyck, S. and G. Peschke, 1995: Grundlagen der hydrologie. Tech. Rep., Verlag für Bauwesen, Berlin. [cited at p. 52]

Feyen, L., J. A. Vrugt, B. O. Nuallain, J. van der Knijff and A. De Roo, 2007: Parameter optimisation and uncertainty assessment for large-scale streamflow simulation with the lisflood model. Journal of Hydrology, **332**, 276–289. [cited at p. 32, 108, 204]

FitzHugh, T. W. and D. S. Mackay, 2000: Impacts of input parameter spatial aggregation on an agricultural nonpoint source pollution model. Journal of Hydrology, **236**, 35–53. [cited at p. 6, 7]

Franchini, M., J. Wendling, C. Obled and E. Todini, 1996: Physical interpretation and sensitivity analysis of the topmodel. Journal of Hydrology, **175**, 293–338. [cited at p. 6, 110]

Gan, T. Y., E. M. Dlamini and G. F. Biftu, 1997: Effects of model complexity and structure, data quality, and objective functions on hydrologic modeling. Journal of Hydrology, **192**, 81–103. [cited at p. 13]

Goovaerts, P., 1999: Geostatistics in soil science: state-of-the-art and perspectives. Geoderma, **89**, 1–45. [cited at p. 5]

Green, W. H. and G. A. Ampt, 1977: Studies on soil physics: I. the flow of air and water through soils. Journal of Agricultural Sciences, **4**. [cited at p. 49, 52]

Hansjürgens, B. and R. M., 2007: Entwicklung und praxistest des verfahrens basinform zur auswahl von maßnahmen im flussgebietsmanagement nach eu-wrrl pilotgebiet döllnitz . Tech. Rep., Helmholz-Zentrum für Umwelforschung - UFZ. [cited at p. 46]

Helton, J. C. and F. J. Davis, 2003: Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliability Engineering & System Safety, **81**, 23–69. [cited at p. 17, 19, 20]

Heuvelink, G. B. M., 1998: Uncertainty analysis in environmental modelling under a change of spatial scale. Nutrient Cycling in Agroecosystems, **50**, 255–264. [cited at p. 13]

Hofer, E., M. Kloos, B. Krzykacz-Hausmann, J. Peschke and M. Woltereck, 2002: An approximate epistemic uncertainty analysis approach in the presence of epistemic and aleatory uncertainties. Reliability Engineering & System Safety, **77**, 229–238. European Safety and Reliability Conference (ESREL) SEP, 2001 TURIN, ITALY. [cited at p. 37]

Janssen, P. H. M. and P. S. C. Heuberger, 1995: Calibration of process-oriented models. Ecological Modelling, **83**, 55–66. [cited at p. 13]

Kavetski, D., G. Kuczera and S. W. Franks, 2006: Bayesian analysis of input uncertainty in hydrological modeling: 2. application. Water Resources Research, **42**. [cited at p. 37]

Kuczera, G. and E. Parent, 1998: Monte carlo assessment of parameter uncertainty in conceptual catchment models: the metropolis algorithm. Journal of Hydrology, **211**, 69–85. [cited at p. 18, 33, 34]

Legates, D. R. and G. J. McCabe, 1999: Evaluating the use of ‟goodness-of-fit″measures in hydrologic and hydroclimatic model validation. Water Resources Research, **35**, 233–241. [cited at p. 57, 58]

Lemons, J., 1996: Scientific Uncertainty and Environmental Problem Solving. Cambridge, MA (Blackwell Science). [cited at p. 2]

Mantovan, P. and E. Todini, 2006: Hydrological forecasting uncertainty assessment: Incoherence of the glue methodology. Journal of Hydrology, **330**, 368–381. [cited at p. 27]

McMillan, H. and M. Clark, 2009: Rainfall-runoff model calibration using informal likelihood measures within a Markov chain Monte Carlo sampling scheme. Water Resources Research, **45**. [cited at p. 109]

Merz, B. and A. H. Thieken, 2005: Separating natural and epistemic uncertainty in flood frequency analysis. Journal of Hydrology, **309**, 114–132. [cited at p. 5]

Montanari, A., 2005: Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. Water Resources Research, **41**. [cited at p. 27]

Moore, C. and J. Doherty, 2005: Role of the calibration process in reducing model predictive error. Water Resources Research, **41**. [cited at p. 57]

Mowrer, H. T., 2000: Uncertainty in natural resource decision support systems: sources, interpretation, and importance. Computers and Electronics in Agriculture, **27**, 139–154. [cited at p. 2, 6]

Perrin, C., C. Michel and V. Andreassian, 2001: Does a large number of parameters enhance model performance? comparative assessment of common catchment model structures on 429 catchments. Journal of Hydrology, **242**, 275–301. [cited at p. 13]

Peschke, G., 1977: Ein zweistufiges Modell der Infiltration von Regen in geschichtete Böden. Acta hydrophysica, **22(1)**. [cited at p. 49]

Pöhler, H., 2006: Anpassung von WaSiM-ETH und die Erstellung und Berechnung von Landnutzungs- und Klimaszenarien fr die Niederschlag-Abfluss-Modellierung am Beispiel des Osterzgebirges. Ph.D Thesis, TU Freiberg. [cited at p. 50, 52, 53, 54]

Quinn, P., K. Beven, P. Chevallier and O. Planchon, 1991: The prediction of hillslope flow paths for distributed hydrological modeling using digital terrain models. Hydrological Processes, **5**, 59–79. [cited at p. 6]

QUINN, P., K. Beven and R. LAMB, 1995: The ln(a/tan-beta) index - how to calculate it and how to use it within the topmodel framework. Hydrological Processes, **9**, 161–182. [cited at p. 110]

Reckhow, K., 1994: Water-quality simulation modeling and uncertainty analysis for risk assessment and decision-making. Ecological Modelling, **72**, 1–20. [cited at p. 2]

Refsgaard, J., M. Thorsen, J. Jensen, S. Kleeschulte and S. Hansen, 1999: Large scale modelling of groundwater contamination from nitrate leaching. Journal OF Hydrology, **221**, 117–140. [cited at p. 6, 7]

Refsgaard, J. C. and J. Knudsen, 1996: Operational validation and intercomparison of different types of hydrological models. Water Resources Research, **32**, 2189–2202. [cited at p. 14]

Refsgaard, J. C., J. P. van der Sluijs, J. Brown and P. van der Keur, 2006: A framework for dealing with uncertainty due to model structure error. Advances in Water Resources, **29**, 1586–1597. [cited at p. 12]

Richter, D., 1995: Ergebnisse methodischer Untersuchungen zur Korrektur des systematischen Messfehlers des Hellmann-Niederschlagsmessers. Tech. Rep. 194, Berichte des Deutschen Wetterdienstes. [cited at p. 11, 50, 56, 57, 61, 204]

Rode, M., 2008: Analysis and modelling of nutrient transport and transformation processes on the catchment scale. Habilitation, University of Potsdam. [cited at p. 7, 204]

Rode, M. and G. Wenk, 2006: Weiße elster river basin data report. Tech. Rep., UFZ - Helmholtz Centre for Environmental Research. [cited at p. 10, 11]

Rode, M. and G. Wink, 2005: Harmonirib: Guidelines for assessing data uncertainty in river basin management studies. Tech. Rep., HarmoniRiB. [cited at p. 8, 9]

Schulla, J., 1997: Hydrologische modellierung von flussgebieten zur abschaetzung der folgen von klimaaenderungen. Ph.D Thesis, ETH Zurich. [cited at p. 47, 49, 50]

Schulla J., J. K., 1998: Modellbeschreibung WaSiM-ETH, Stand 21. Dezember 1998. [cited at p. 49, 50, 52, 53, 54]

Schulla J., J. K., 2001: Model Description WaSiM-ETH, last updated 19th july 2001. [cited at p. 49, 50, 52, 53, 54, 56]

Skahill, B. E. and J. Doherty, 2006: Efficient accommodation of local minima in watershed model calibration. Journal of Hydrology, **329**, 122–139. [cited at p. 54]

Snowling, S. D. and J. R. Kramer, 2001: Evaluating modelling uncertainty for model selection. Ecological Modelling, **138**, 17–30. [cited at p. 13]

Spear, R., T. Grieb and N. Shang, 1994: Parameter uncertainty and interaction in complex environmental-models. Water Resources Research, **30**, 3159–3169. [cited at p. 18, 204]

Stedinger, J. R., R. M. Vogel, S. U. Lee and R. Batchelder, 2008: Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resources Research, **44**. [cited at p. 109]

Van Rompaey, A. J. J. and G. Govers, 2002: Data quality and model complexity for regional scale soil erosion prediction. International Journal of Geographical Information Science, **16**, 663–680. [cited at p. 13, 14]

VanderPerk, M., 1997: Effect of model structure on the accuracy and uncertainty of results from water quality models. Hydrological Processes, **11**, 227–239. [cited at p. 14]

Vazquez, R. F., L. Feyen, J. Feyen and J. C. Refsgaard, 2002: Effect of grid size on effective parameters and model performance of the mike-she code. Hydrological Processes, **16**, 355–372. [cited at p. 7]

Watson, F. G. R., R. B. Grayson, R. A. Vertessy and T. A. McMahon, 1998: Large-scale distribution modelling and the utility of detailed ground data. Hydrological Processes, **12**, 873–888. [cited at p. 5]

de Wit, M. J. M. and E. J. Pebesma, 2001: Nutrient fluxes at the river basin scale. ii: the balance between data availability and model complexity. Hydrological Processes, **15**, 761–775. [cited at p. 14]

Wolock, D. and C. Price, 1994: Effects of digital elevation model map scale and data resolution on a topography-based watershed model. Water Resources Research, **30**, 3041–3052. [cited at p. 6]

Zhang, W. and D. Montgomery, 1994: Digital elevation model grid size, landscape representation, and hydrologic simulations. Water Resources Research, **30**, 1019–1028. [cited at p. 6]

Zhihua, R. and L. Mingelin, 2007: Errors and correction of precipitation measurements in China. Advances in Atmppheric Sciences, **24**, 449–458. [cited at p. 61]

Appendix A

Weiße Elster Figures

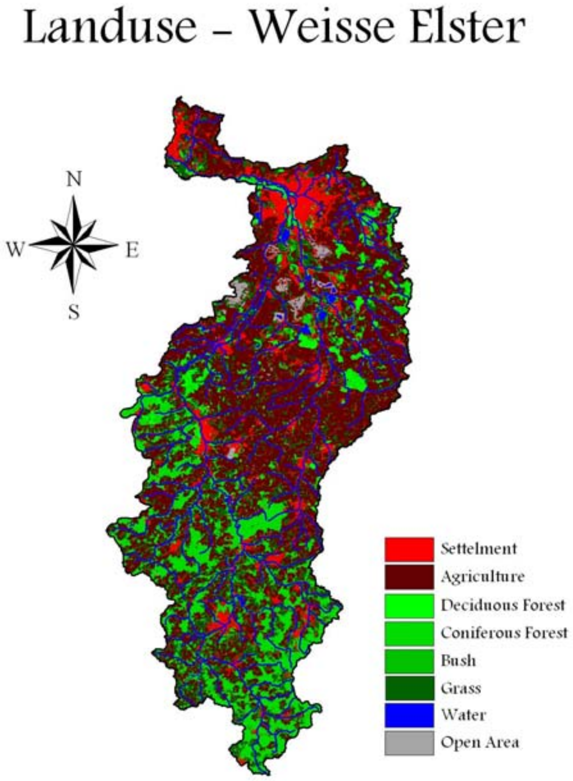


Figure A.1: The land use in the Weiße Elster catchment.

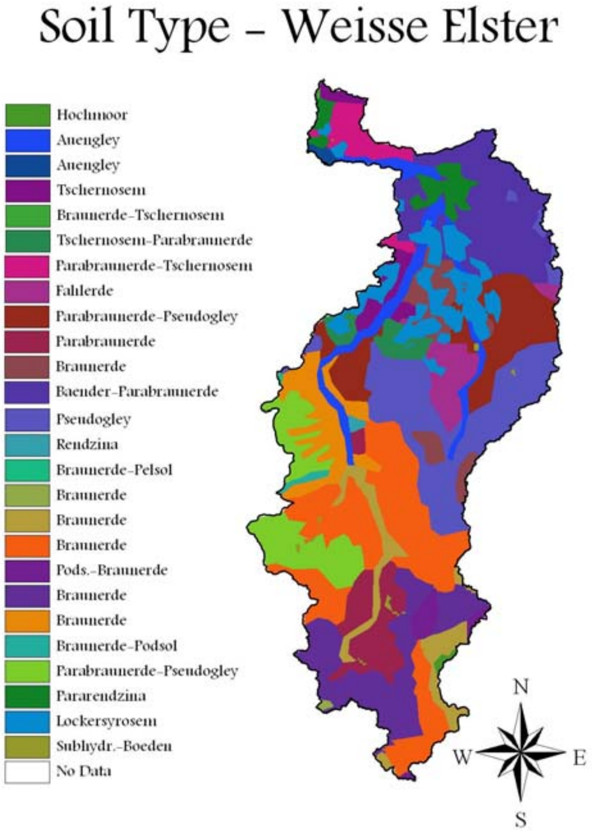


Figure A.2: The soil map for the Weiße Elster catchment.

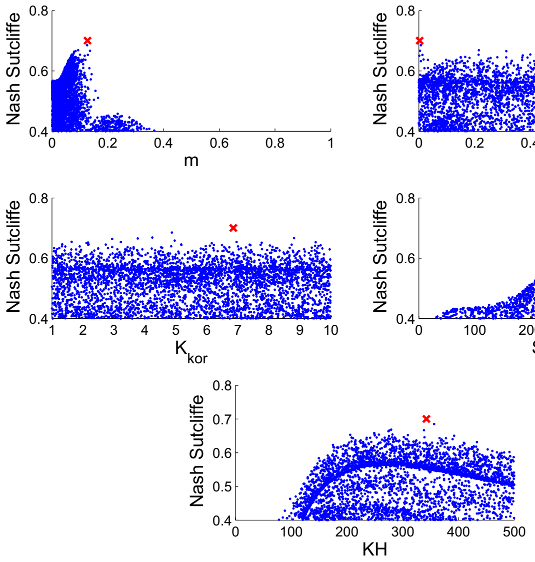


Figure A.3: The dotty plots for the Gera-Langenberg catchment using a 4000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

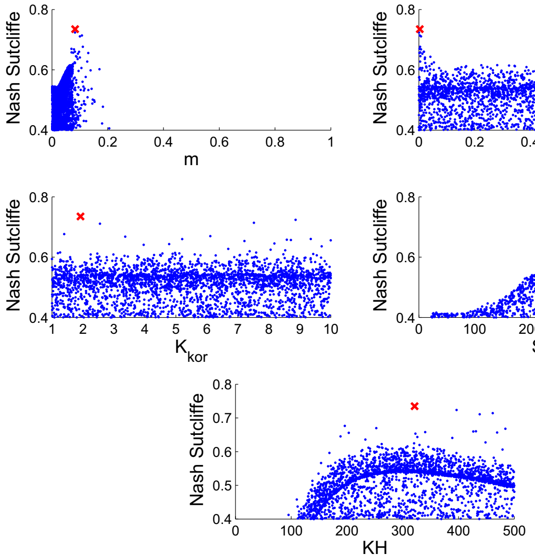


Figure A.4: The dotty plots for the Gera-Langenberg catchment using a 2000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

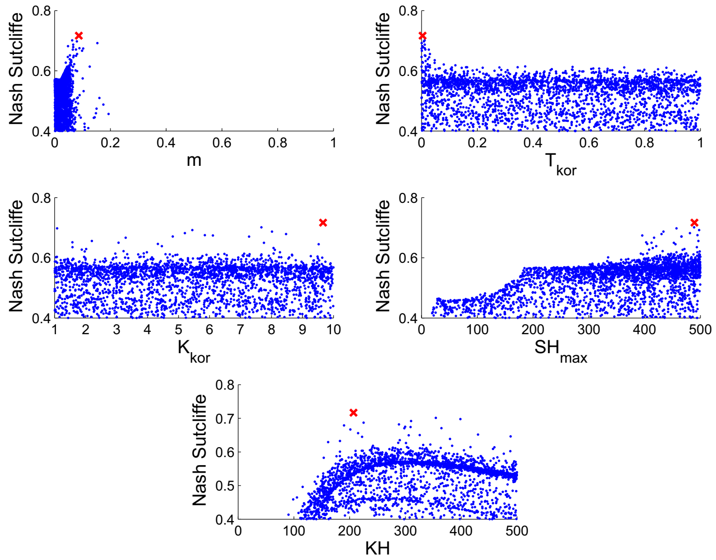


Figure A.5: The dotty plots for the Gera-Langenberg catchment using a 1000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

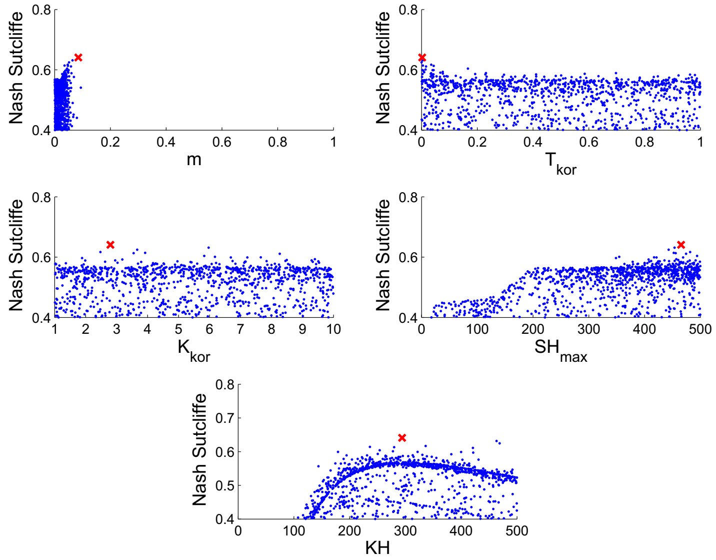


Figure A.6: The dotty plots for the Gera-Langenberg catchment using a 500 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

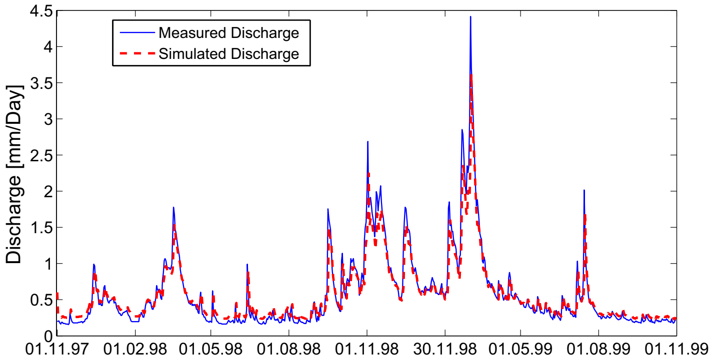


Figure A.7: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 6000 m grid resolution.

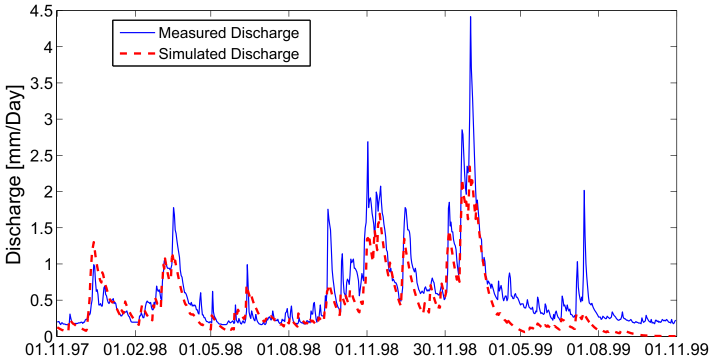


Figure A.8: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 4000 m grid resolution.

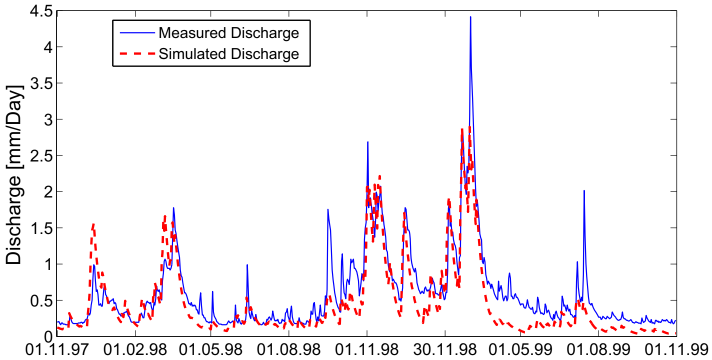


Figure A.9: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 1000 m grid resolution.

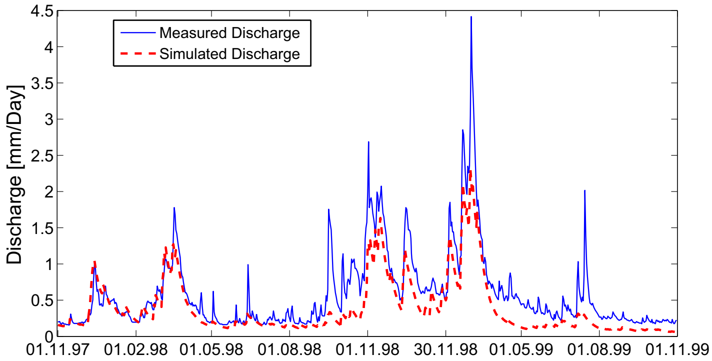


Figure A.10: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 500 m grid resolution.

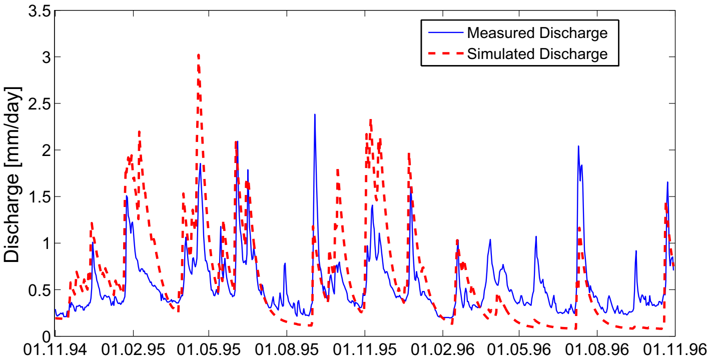


Figure A.11: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 6000 m grid resolution in the validation period.

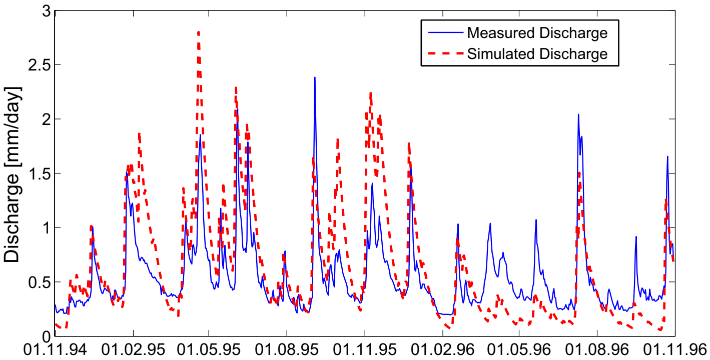


Figure A.12: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 4000 m grid resolution in the validation period.

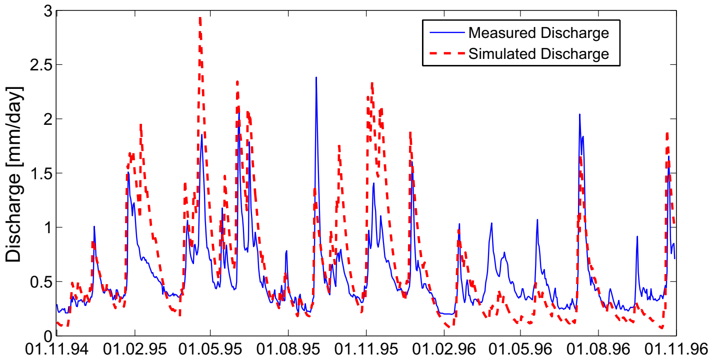


Figure A.13: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 2000 m grid resolution in the validation period.

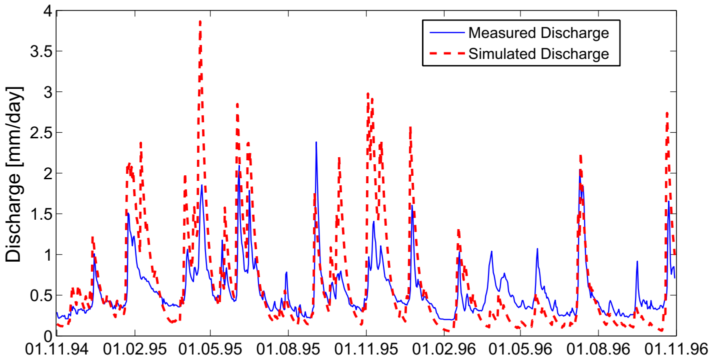


Figure A.14: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 1000 m grid resolution in the validation period.

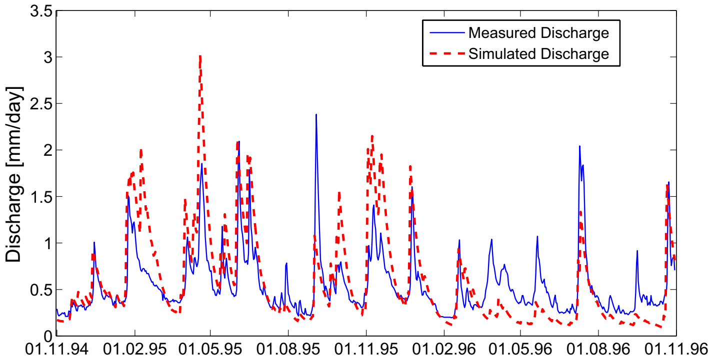


Figure A.15: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from GLUE and a 500 m grid resolution in the validation period.

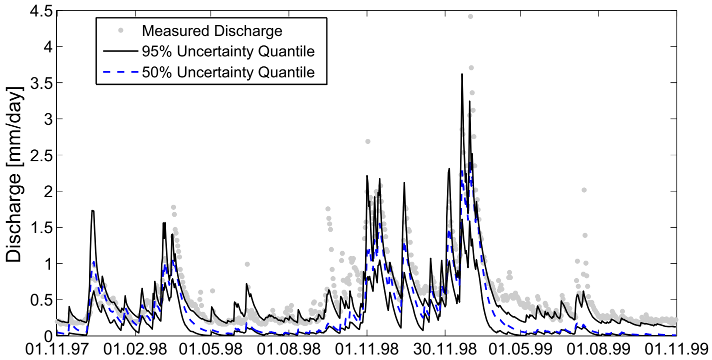


Figure A.16: The 95% and 50% uncertainty quantile for the Gera-Langenberg catchment using the GLUE method and a 4000 *m* grid resolution.

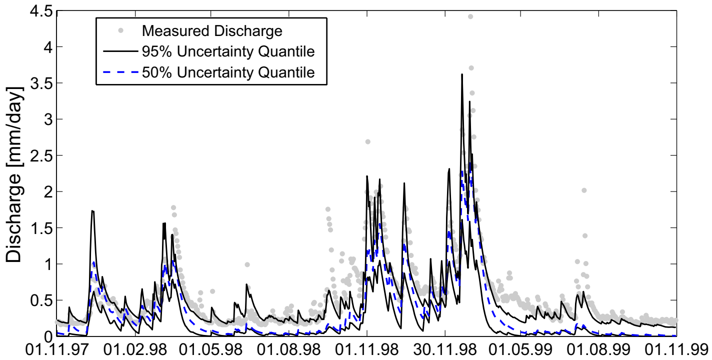


Figure A.17: The 95% and 50% uncertainty quantile for the Gera-Langenberg catchment using the GLUE method and a 2000 *m* grid resolution.

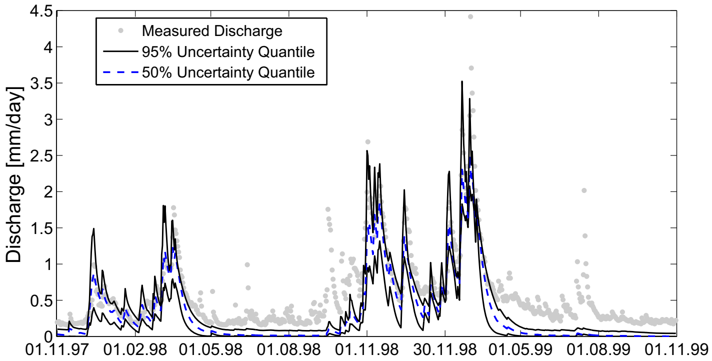


Figure A.18: The 95% and 50% uncertainty quantile for the Gera-Langenberg catchment using the GLUE method and a 1000 *m* grid resolution.

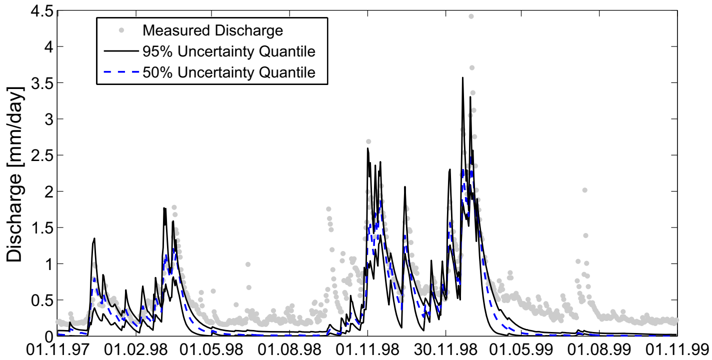


Figure A.19: The 95% and 50% uncertainty quantile for the Gera-Langenberg catchment using the GLUE method and a 500 m grid resolution.

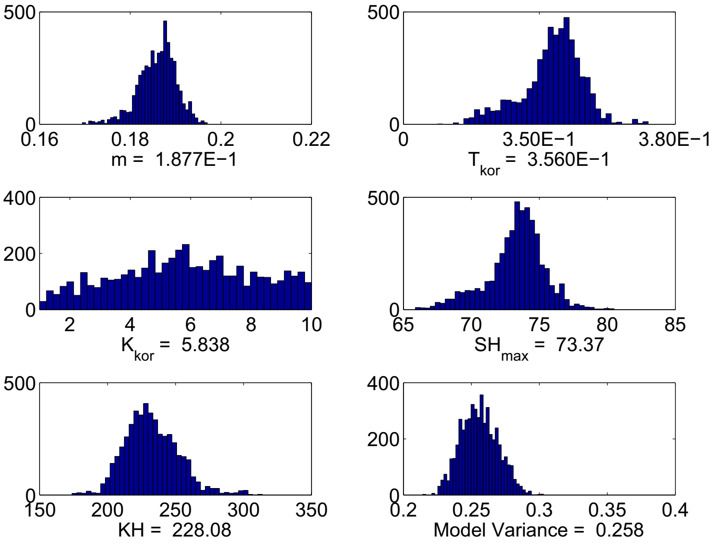


Figure A.20: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method without autocorrelation and the 6000 m grid resolution.

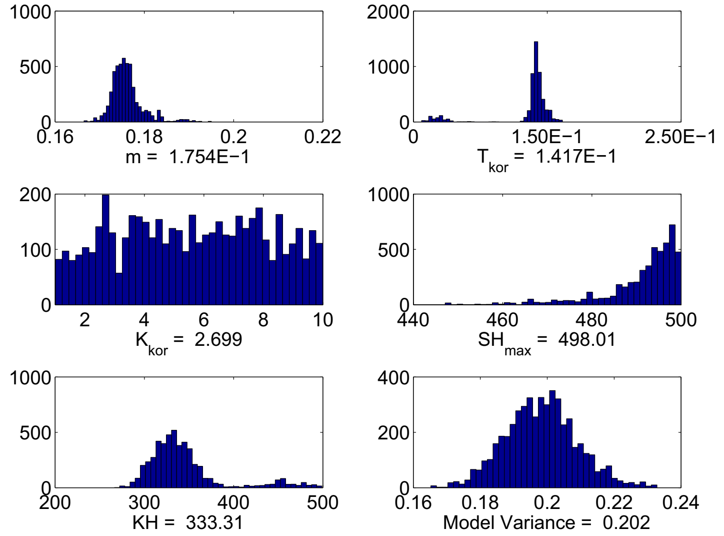


Figure A.21: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method without autocorrelation and the 4000 m grid resolution.

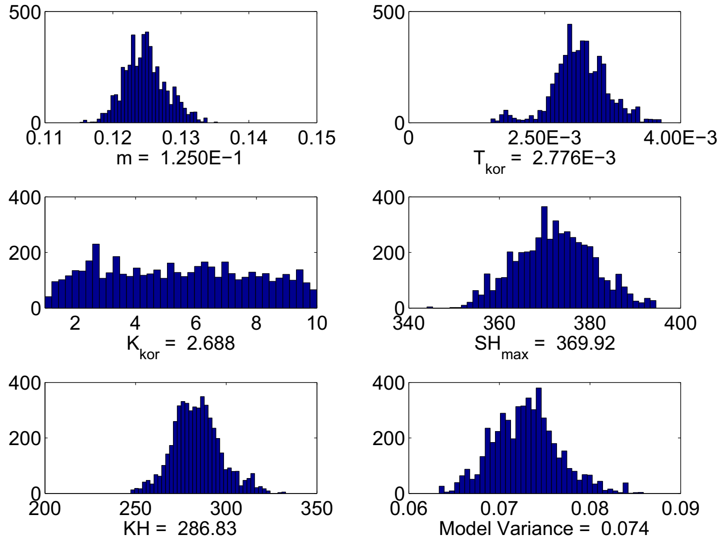


Figure A.22: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method without autocorrelation and the 1000 m grid resolution.

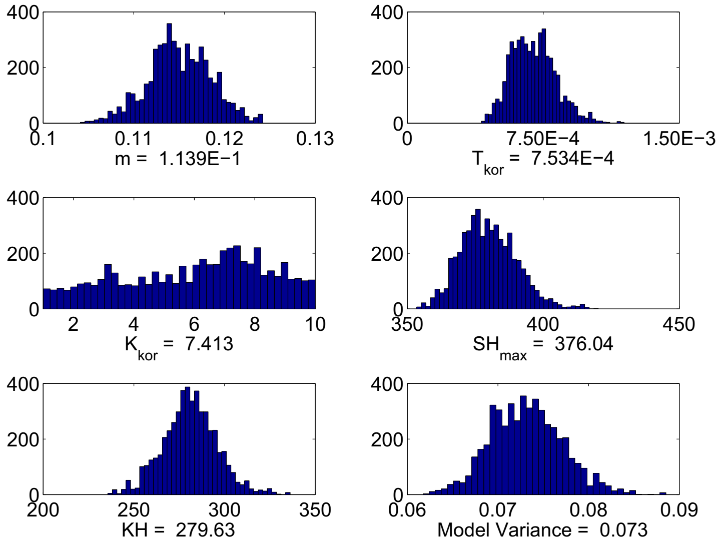


Figure A.23: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method without autocorrelation and the 500 m grid resolution.

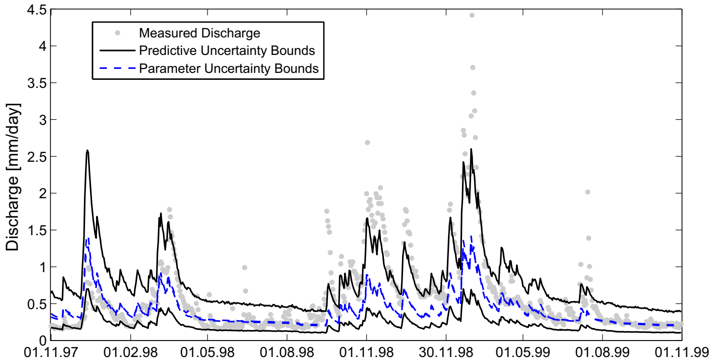


Figure A.24: The predictive and parameter uncertainty bounds for the Gera-Langenberg catchment using MC^2 without autocorrelation and a 6000 m grid resolution.

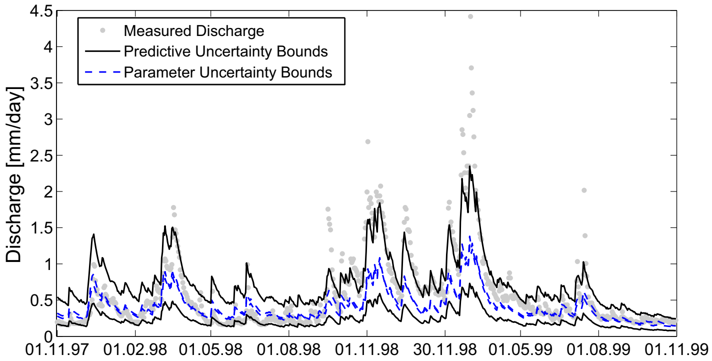


Figure A.25: The predictive and parameter uncertainty bounds for the Gera-Langenberg catchment using MC^2 without autocorrelation and a 4000 m grid resolution.

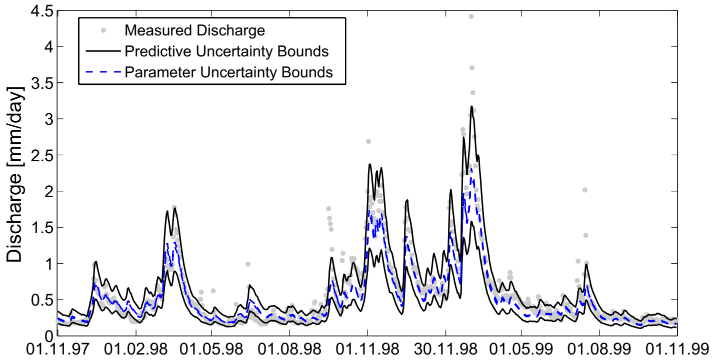


Figure A.26: The predictive and parameter uncertainty bounds for the Gera-Langenberg catchment using MC^2 without autocorrelation and a 1000 m grid resolution.

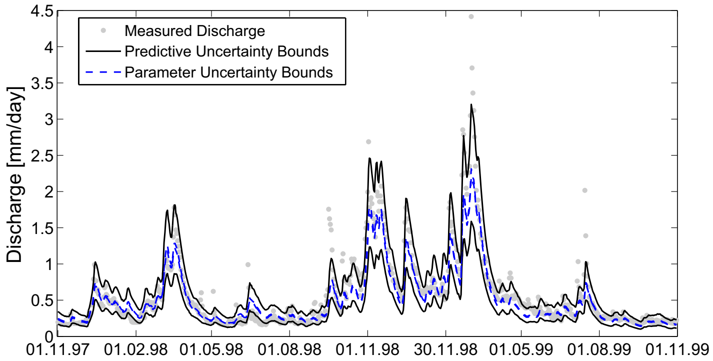


Figure A.27: The predictive and parameter uncertainty bounds for the Gera-Langenberg catchment using MC^2 without autocorrelation and a 500 m grid resolution.

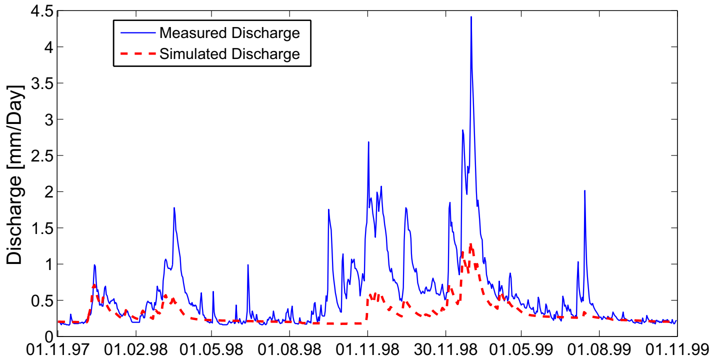


Figure A.28: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 6000 m grid resolution.

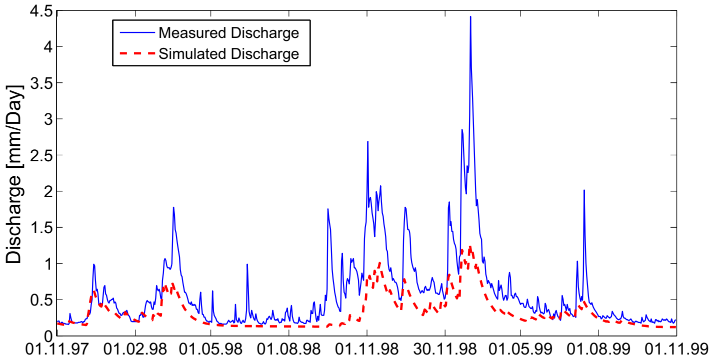


Figure A.29: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 4000 m grid resolution.

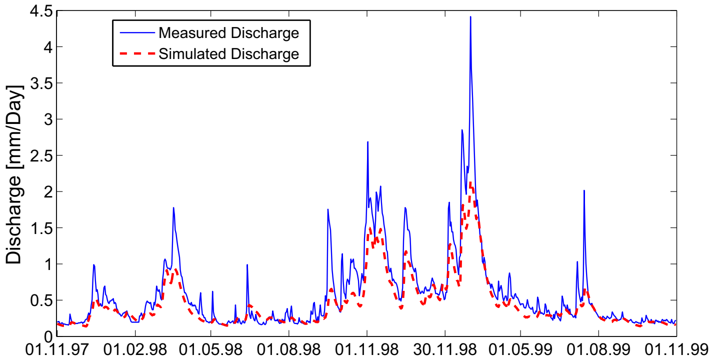


Figure A.30: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 1000 m grid resolution.

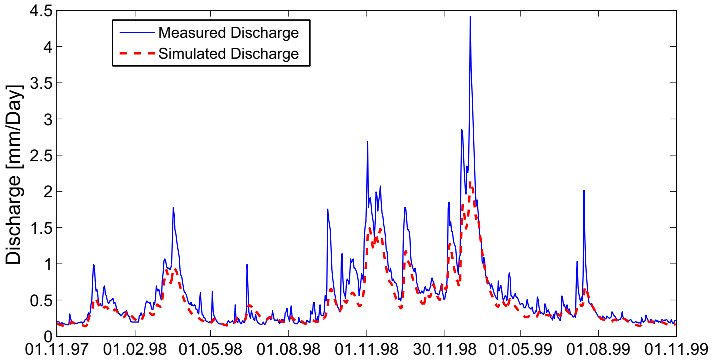


Figure A.31: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 500 m grid resolution.

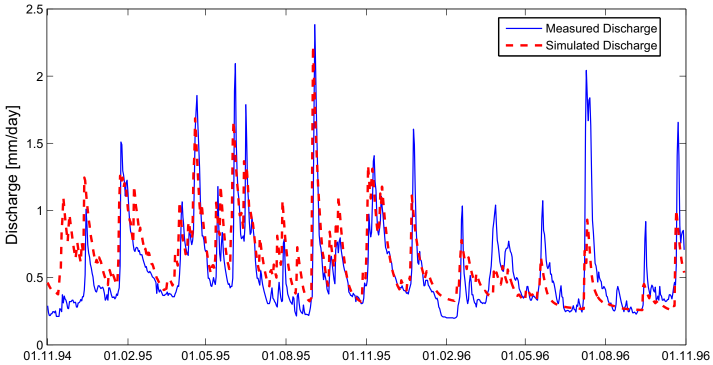


Figure A.32: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 6000 m grid resolution in the validation period.

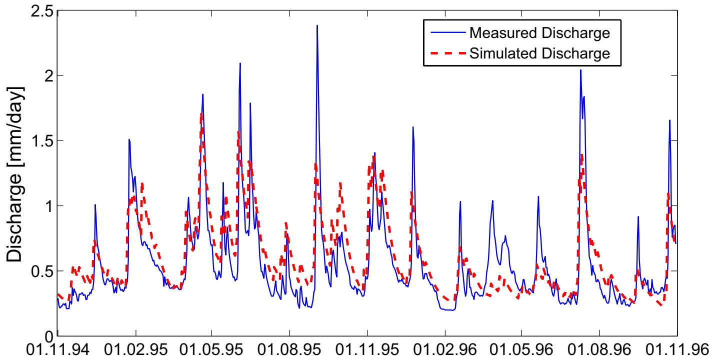


Figure A.33: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 4000 m grid resolution in the validation period.

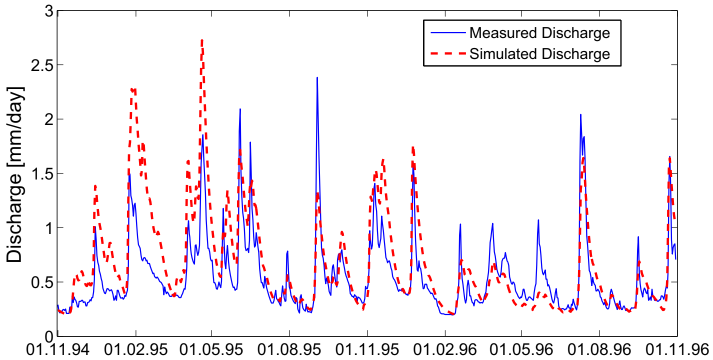


Figure A.34: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 2000 m grid resolution in the validation period.

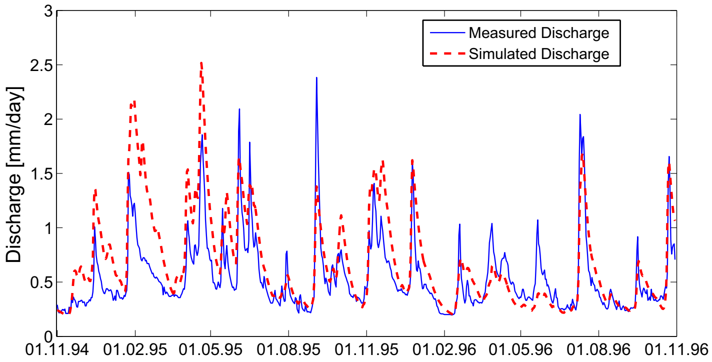


Figure A.35: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 1000 m grid resolution in the validation period.

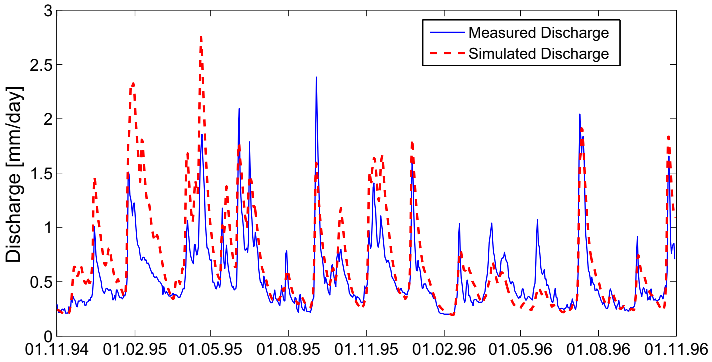


Figure A.36: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 500 m grid resolution in the validation period.

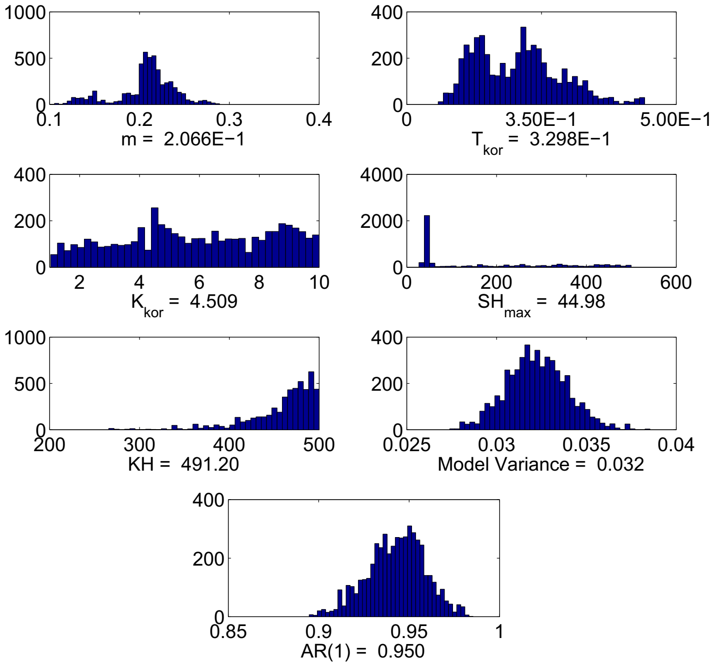


Figure A.37: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method with autocorrelation and the 6000 m grid resolution.

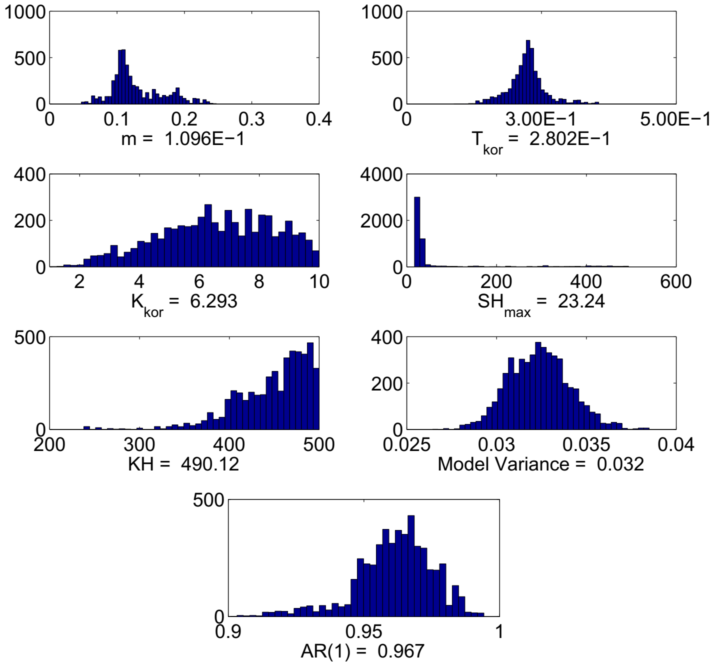


Figure A.38: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method with autocorrelation and the 4000 m grid resolution.

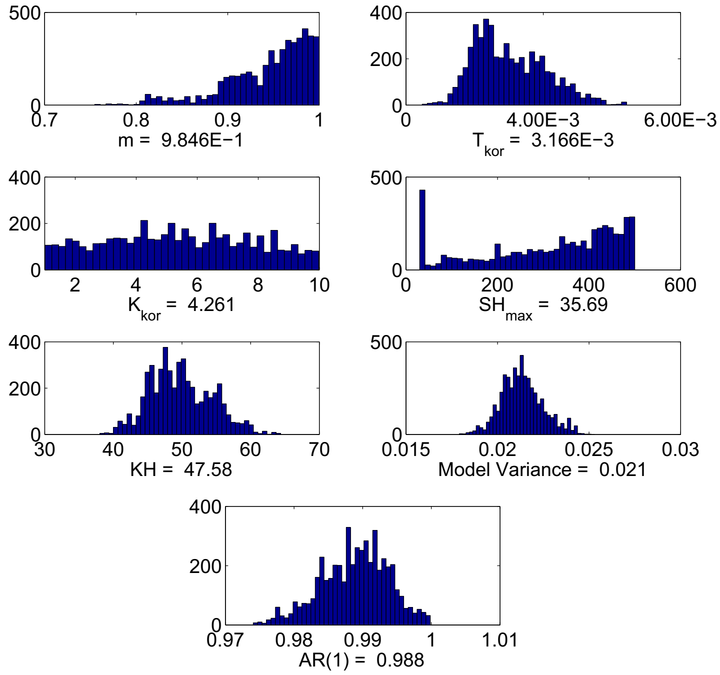


Figure A.39: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method with autocorrelation and the 1000 m grid resolution.

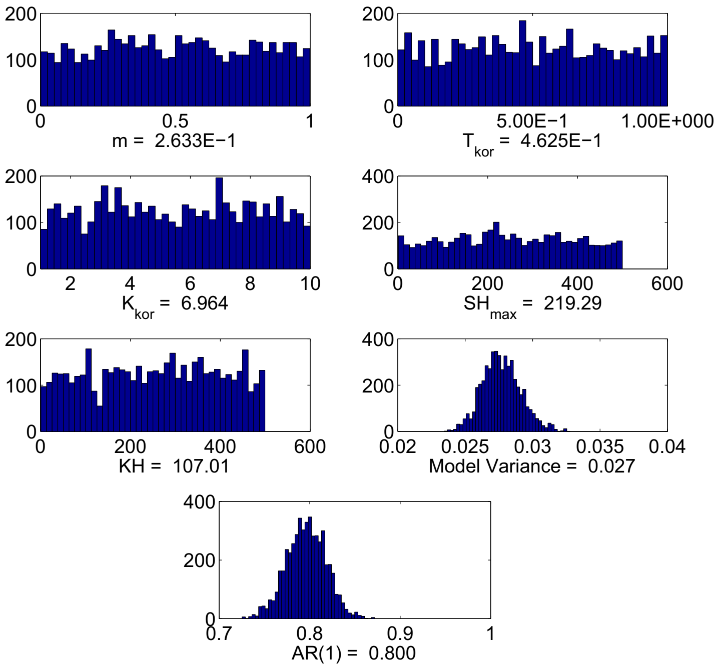


Figure A.40: The parameter posterior distribution for the Gera-Langenberg catchment using the MC^2 method with autocorrelation and the 500 m grid resolution.

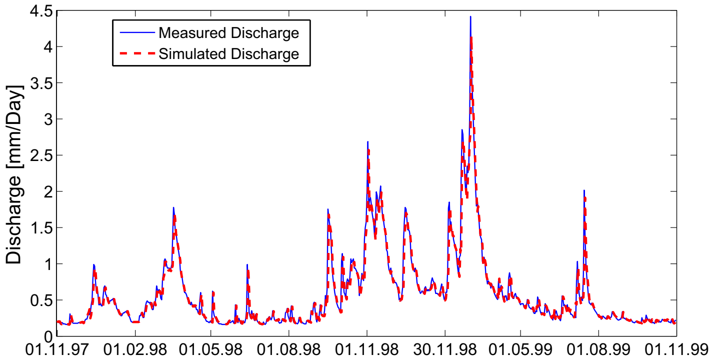


Figure A.41: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 6000 m grid resolution.

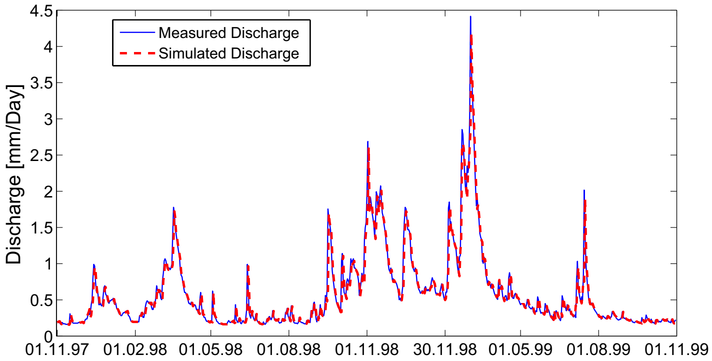


Figure A.42: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 4000 m grid resolution.

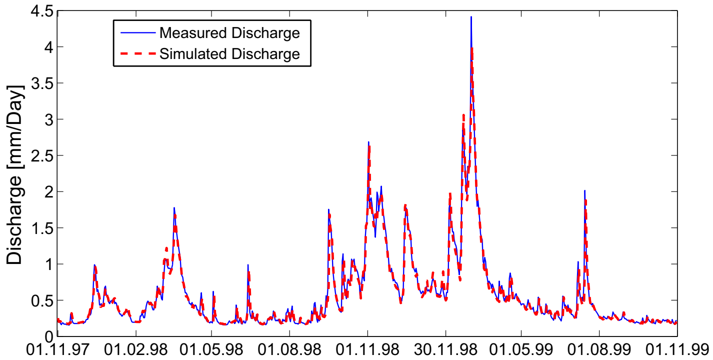


Figure A.43: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 1000 m grid resolution.

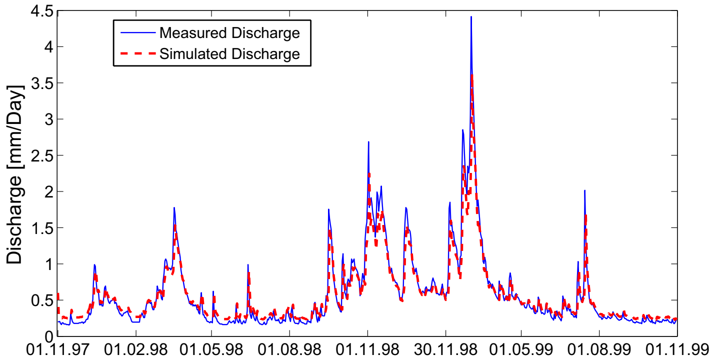


Figure A.44: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 500 m grid resolution.

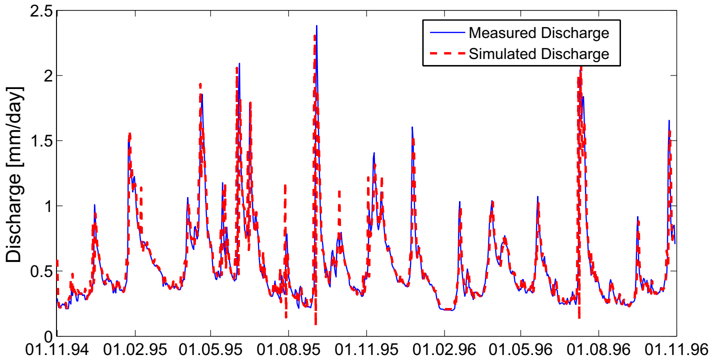


Figure A.45: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 6000 m grid resolution in the validation period.

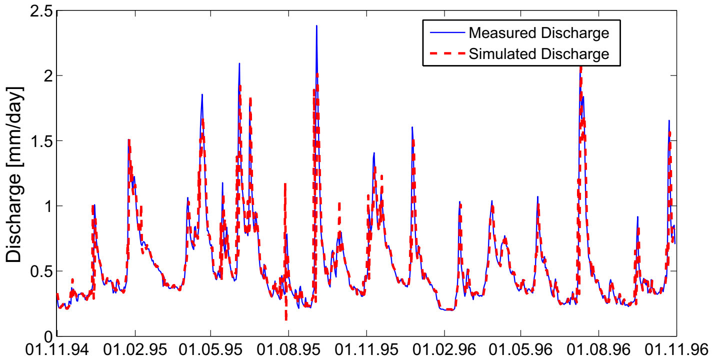


Figure A.46: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 4000 m grid resolution in the validation period.

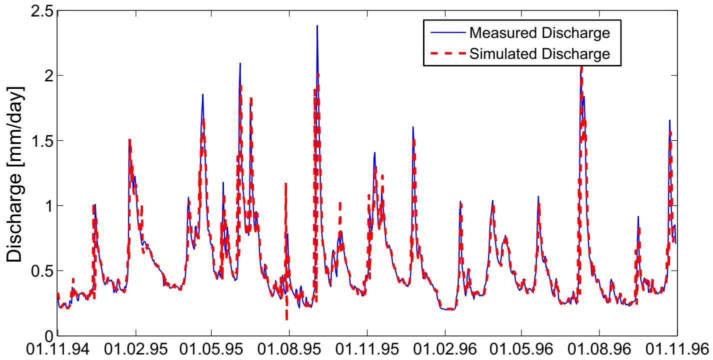


Figure A.47: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 2000 m grid resolution in the validation period.

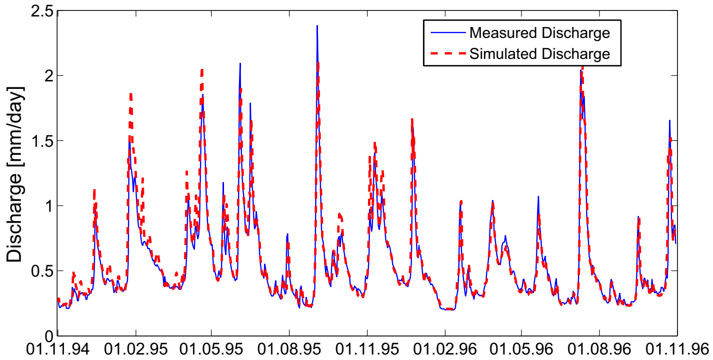


Figure A.48: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 1000 m grid resolution in the validation period.

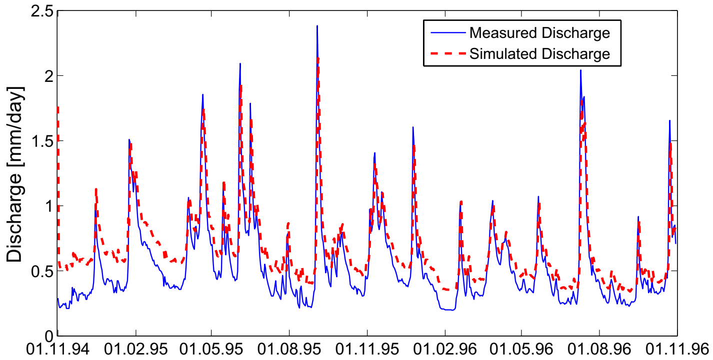


Figure A.49: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from MC^2 and a 500 m grid resolution in the validation period.

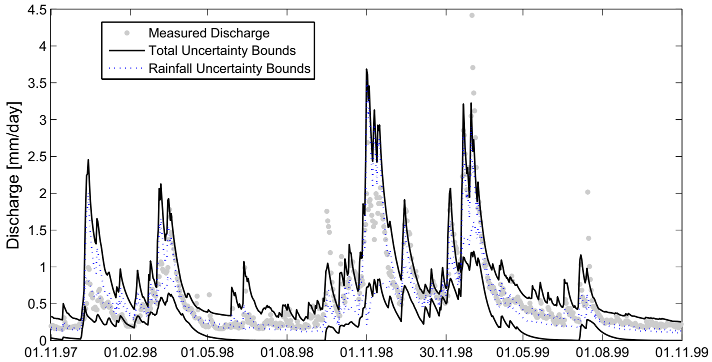


Figure A.50: The total and the rainfall uncertainty bounds for the Gera-Langenberg catchment using the INPUT concept and a 6000 m grid resolution.

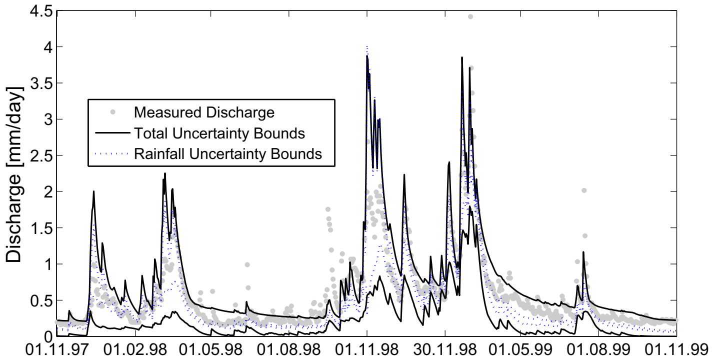


Figure A.51: The total and the rainfall uncertainty bounds for the Gera-Langenberg catchment using the INPUT concept and a 4000 m grid resolution.

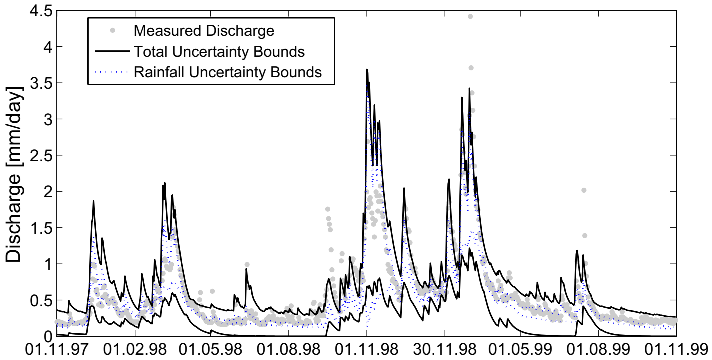


Figure A.52: The total and the rainfall uncertainty bounds for the Gera-Langenberg catchment using the INPUT concept and a 1000 m grid resolution.

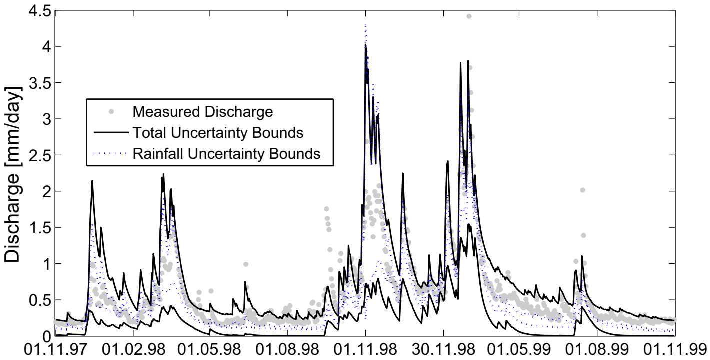


Figure A.53: The total and the rainfall uncertainty bounds for the Gera-Langenberg catchment using the INPUT concept and a 500 m grid resolution.

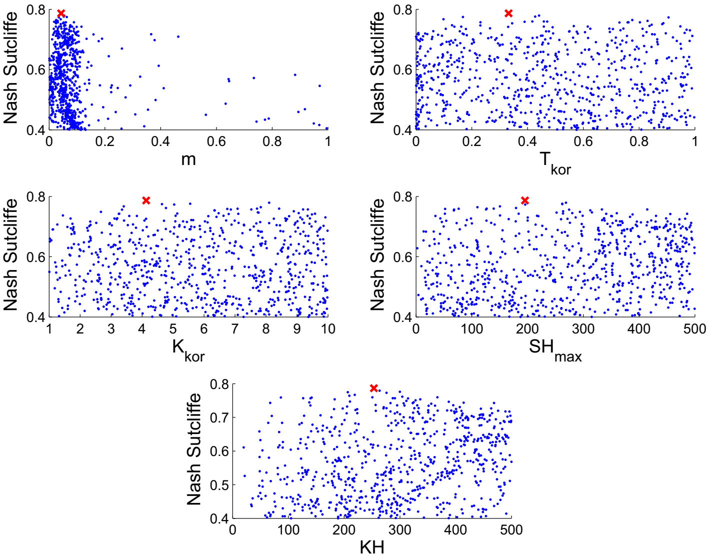


Figure A.54: The dotty plots for the Gera-Langenberg catchment using the INPUT concept and a 6000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

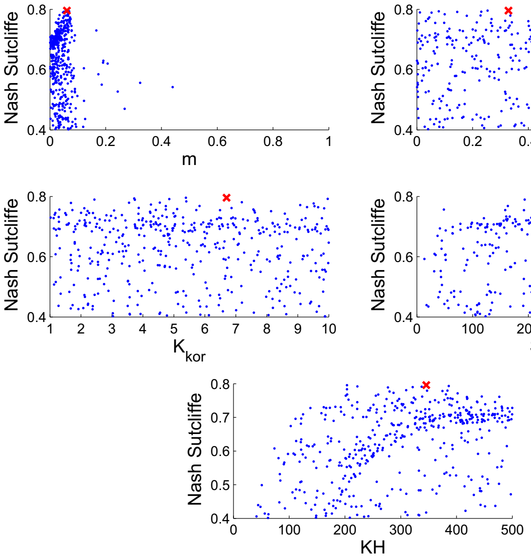


Figure A.55: The dotty plots for the Gera-Langenberg catchment using the INPUT concept and a 4000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

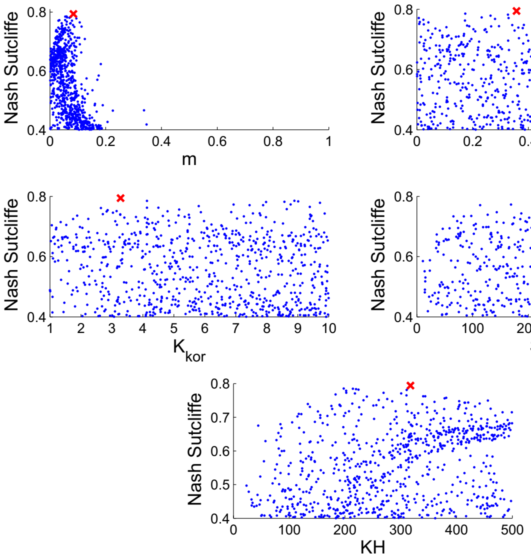


Figure A.56: The dotty plots for the Gera-Langenberg catchment using the INPUT concept and a 1000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

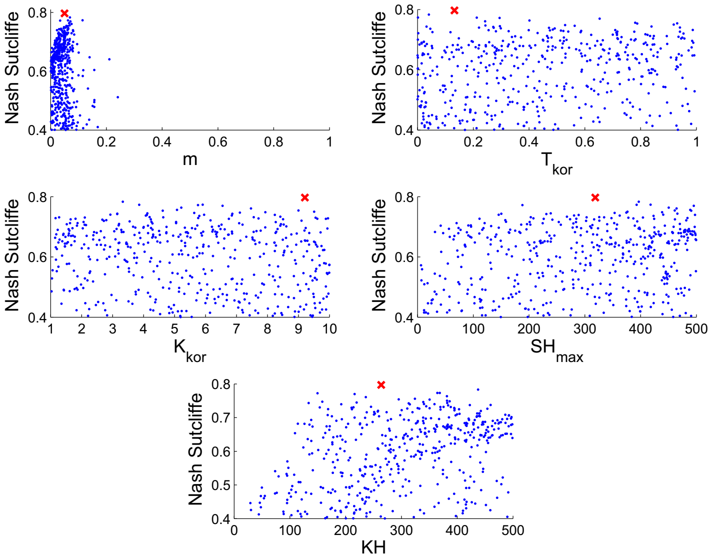


Figure A.57: The dotty plots for the Gera-Langenberg catchment using the INPUT concept and a 500 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

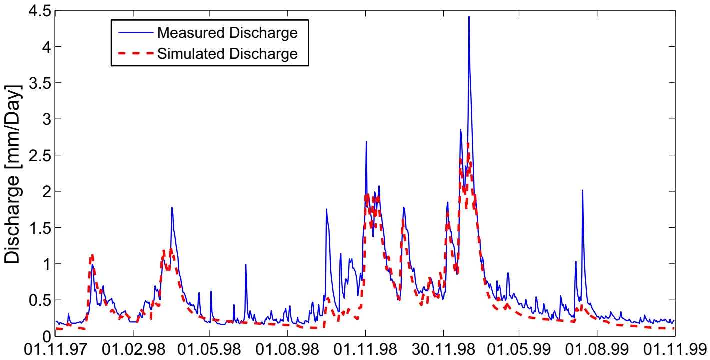


Figure A.58: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 6000 m grid resolution.

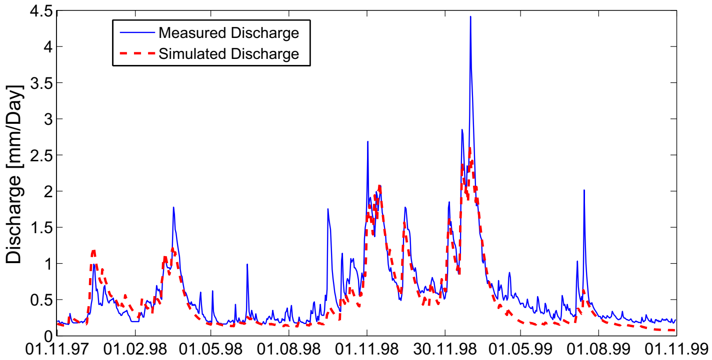


Figure A.59: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 4000 m grid resolution.

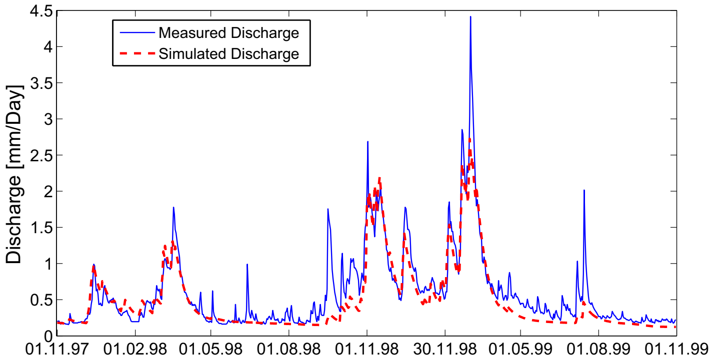


Figure A.60: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 1000 m grid resolution.

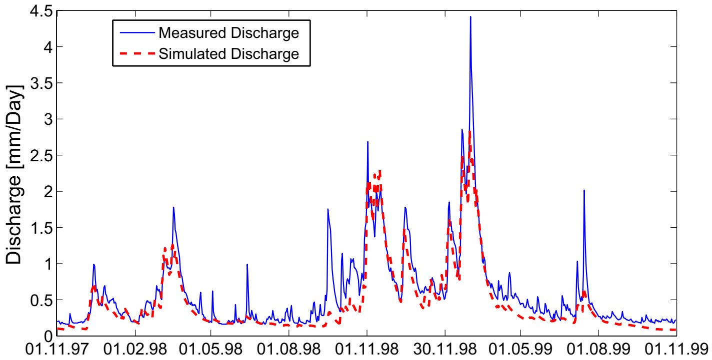


Figure A.61: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 500 m grid resolution.

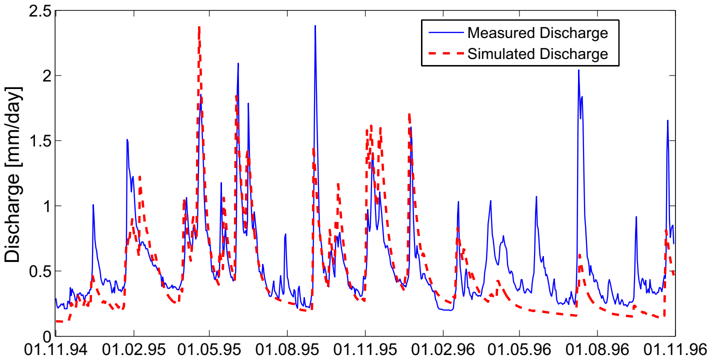


Figure A.62: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 6000 m grid resolution in the validation period.

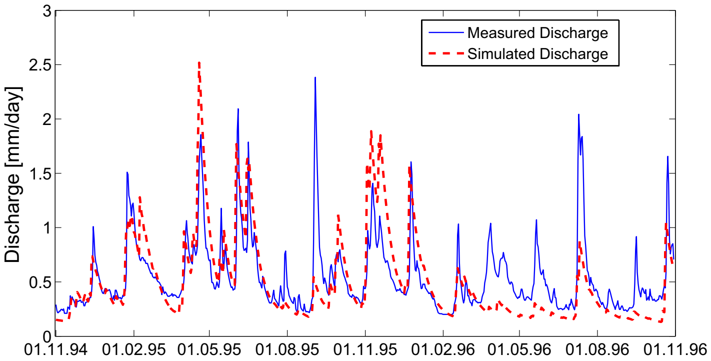


Figure A.63: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 4000 m grid resolution in the validation period.

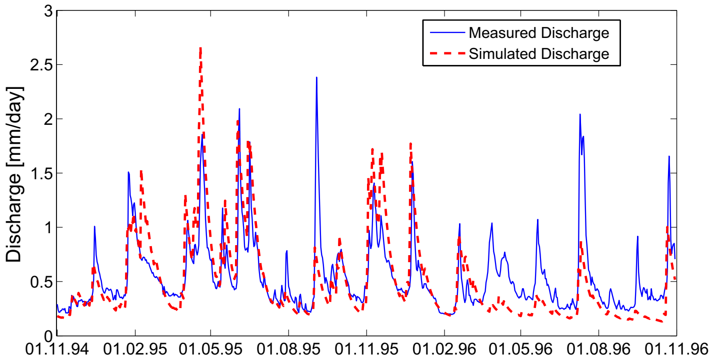


Figure A.64: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 2000 m grid resolution in the validation period.

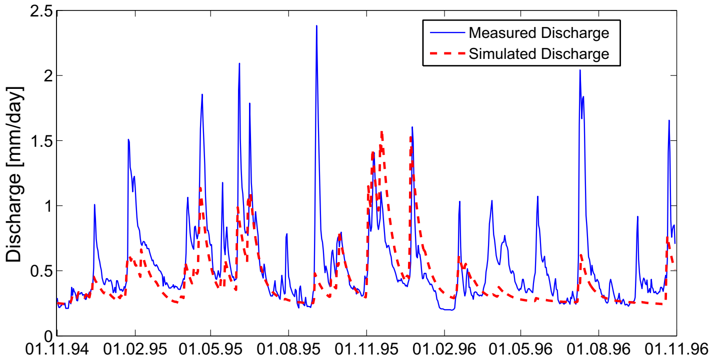


Figure A.65: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 1000 m grid resolution in the validation period.

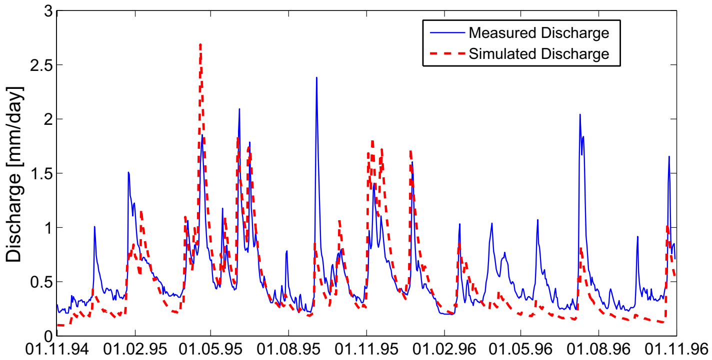


Figure A.66: The measured and simulated discharge for the Gera-Langenberg catchment using the best parameter set from the INPUT concept and a 500 m grid resolution in the validation period.

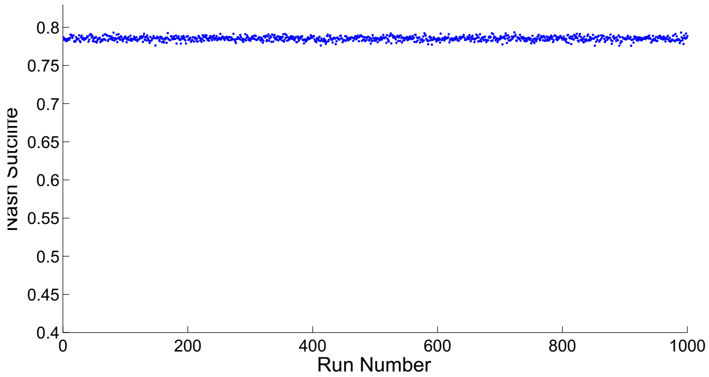


Figure A.67: The Nash Sutcliffe Criteria of the second Run for the Gera-Langenberg catchment unsing using the INPUT concept and a 6000 m grid resolution.

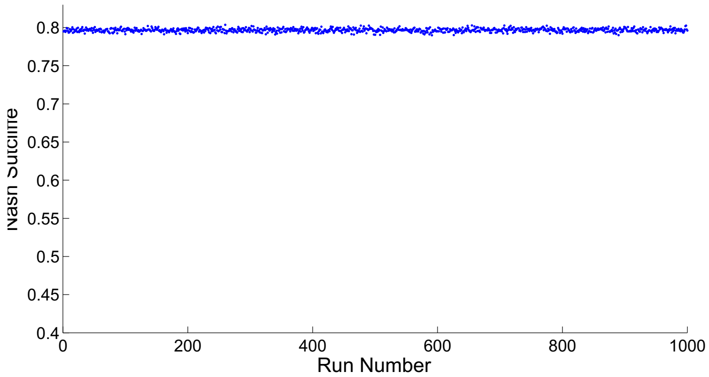


Figure A.68: The Nash Sutcliffe Criteria of the second Run for the Gera-Langenberg catchment unsing using the INPUT concept and a 4000 m grid resolution.

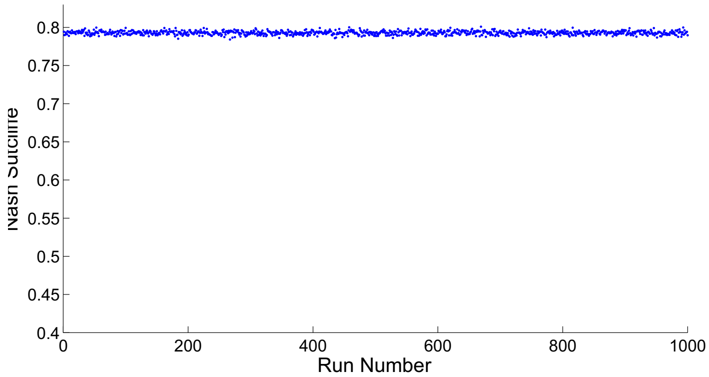


Figure A.69: The Nash Sutcliffe Criteria of the second Run for the Gera-Langenberg catchment unsing using the INPUT concept and a 1000 m grid resolution.

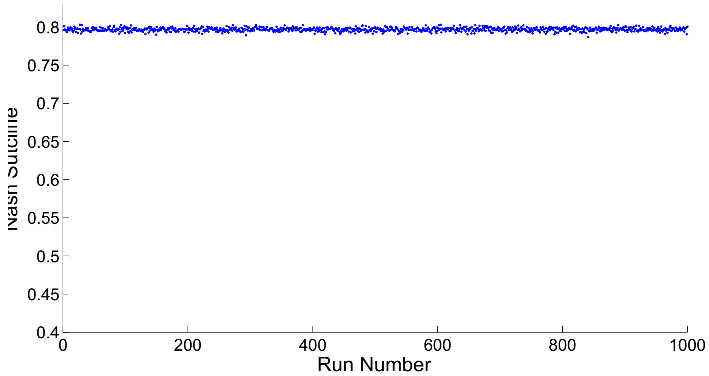


Figure A.70: The Nash Sutcliffe Criteria of the second Run for the Gera-Langenberg catchment unsing using the INPUT concept and a 500 m grid resolution.

Appendix B

Weida Catchment Figures

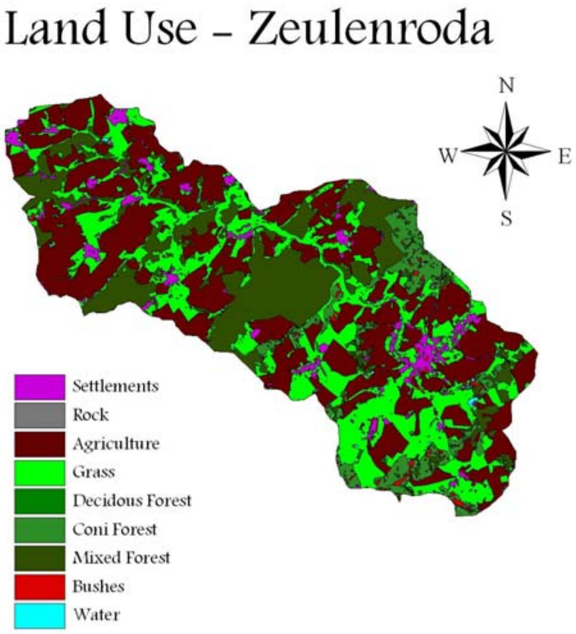


Figure B.1: The land use in the Weida catchment.

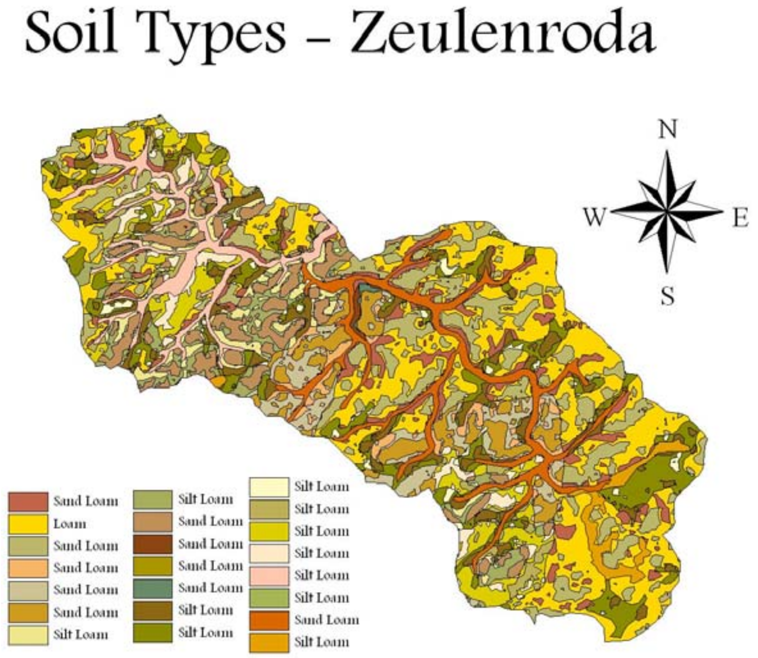


Figure B.2: The soil map for the Weida catchment.

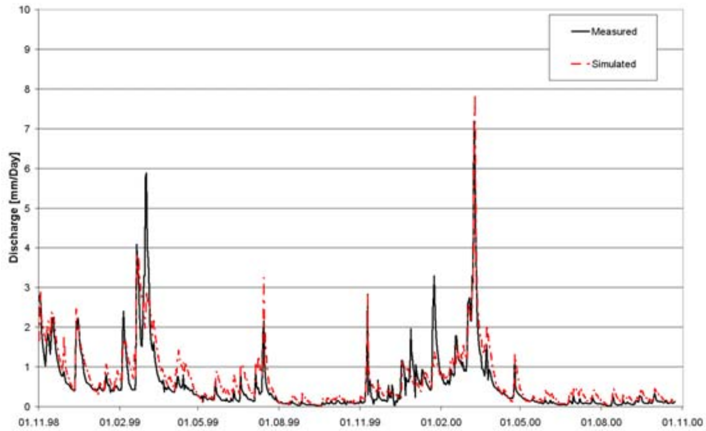


Figure B.3: The measured and simulated discharge for the Weida catchment using the best parameter set from PEST and a 100 m grid resolution.

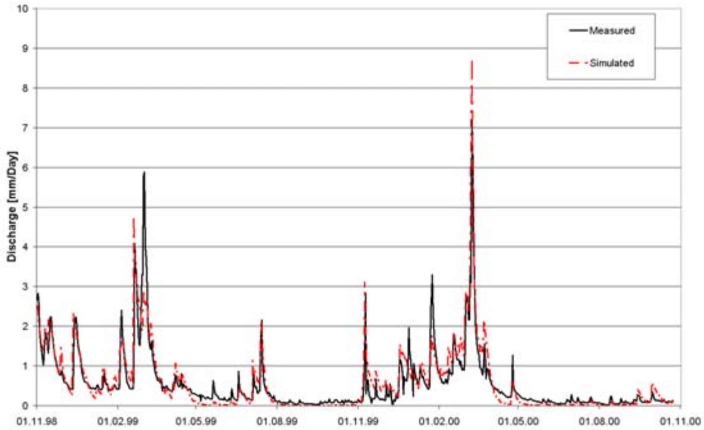


Figure B.4: The measured and simulated discharge for the Weida catchment using the best parameter set from PEST and a 300 m grid resolution.

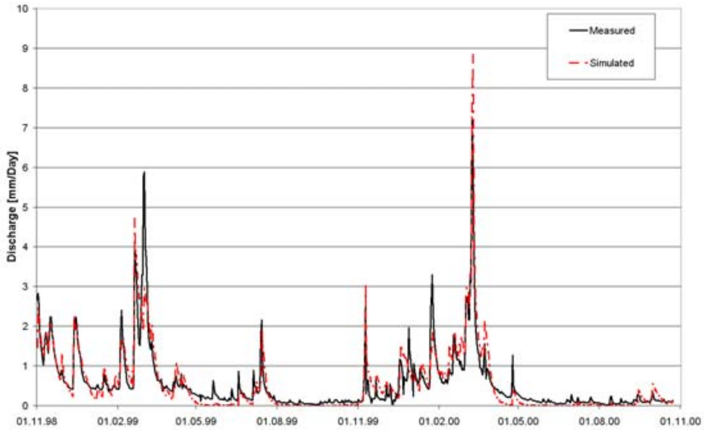


Figure B.5: The measured and simulated discharge for the Weida catchment using the best parameter set from PEST and a 500 m grid resolution.

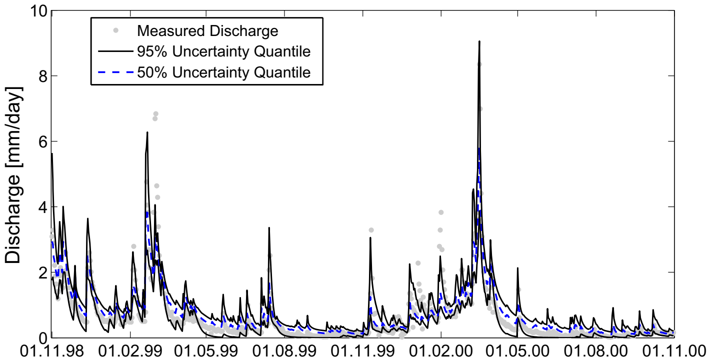


Figure B.6: The 95% and 50% uncertainty quantile for the Weida catchment using the GLUE method and a 100 m grid resolution.

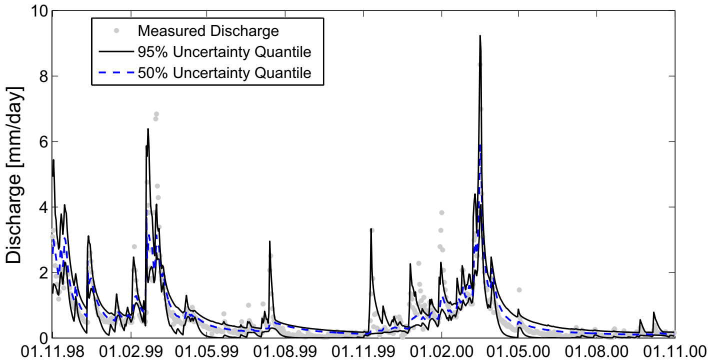


Figure B.7: The 95% and 50% uncertainty quantile for the Weida catchment using the GLUE method and a 300 m grid resolution.

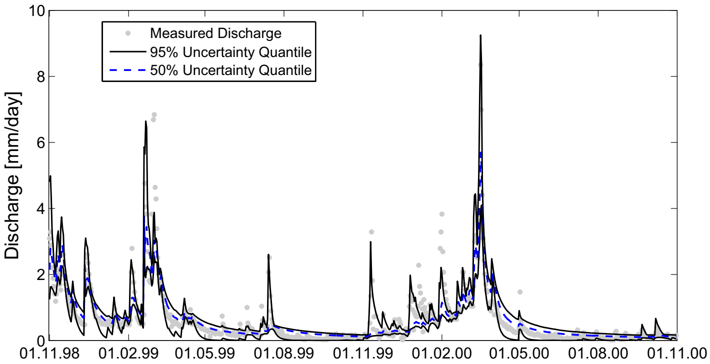


Figure B.8: The 95% and 50% uncertainty quantile for the Weida catchment using the GLUE method and a 500 m grid resolution.

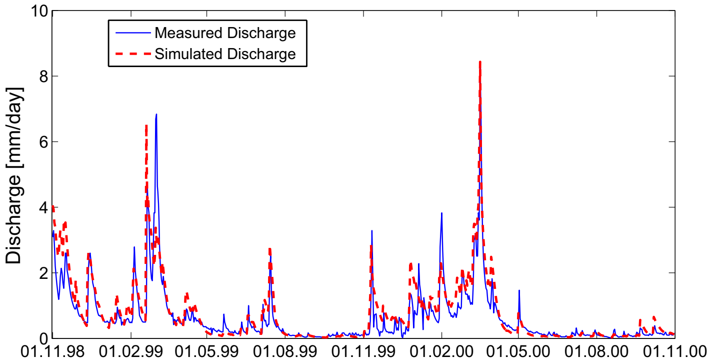


Figure B.9: The measured and simulated discharge for the Weida catchment using the best paramter set from GLUE and a 100 m grid resolution.

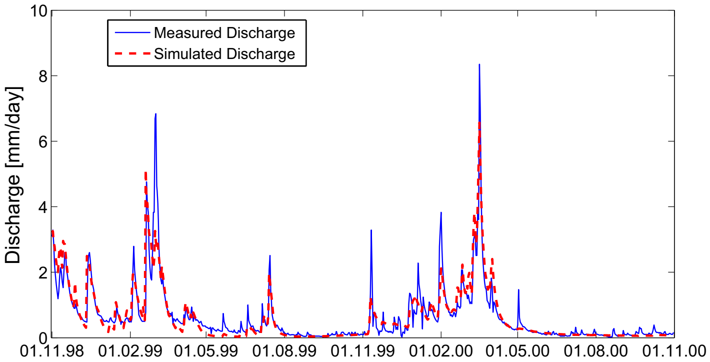


Figure B.10: The measured and simulated discharge for the Weida catchment using the best paramter set from GLUE and a 300 m grid resolution.

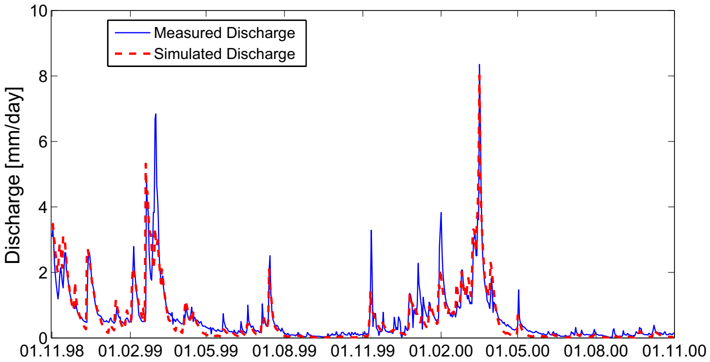


Figure B.11: The measured and simulated discharge for the Weida catchment using the best paramter set from GLUE and a 500 m grid resolution.

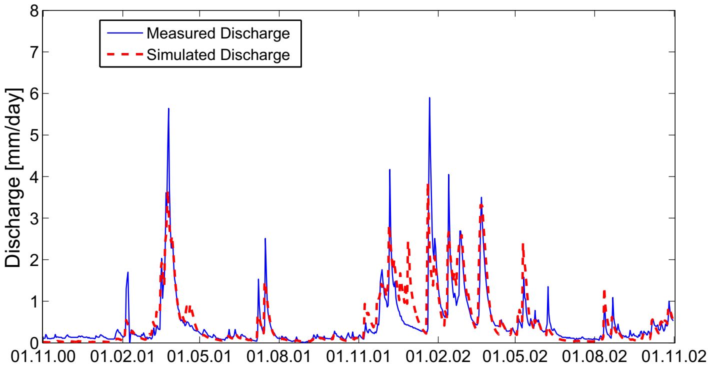


Figure B.12: The measured and simulated discharge for the Weida catchment using the best paramter set from GLUE and a 100 m grid resolution for the validation period.

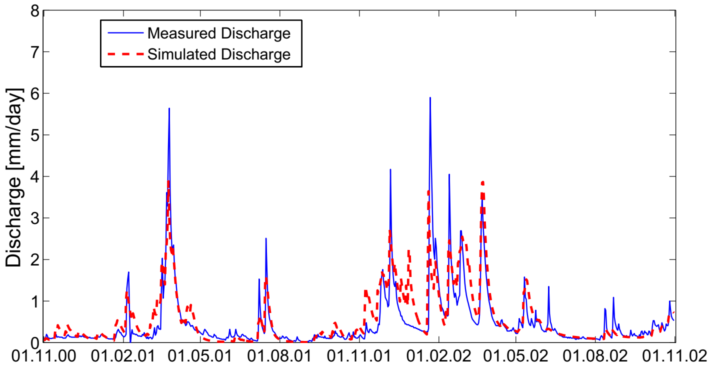


Figure B.13: The measured and simulated discharge for the Weida catchment using the best paramter set from GLUE and a 300 m grid resolution for the validation period.

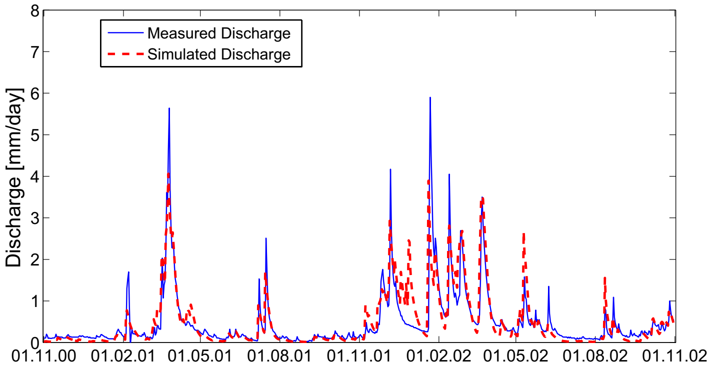


Figure B.14: The measured and simulated discharge for the Weida catchment using the best paramter set from GLUE and a 500 m grid resolution for the validation period.

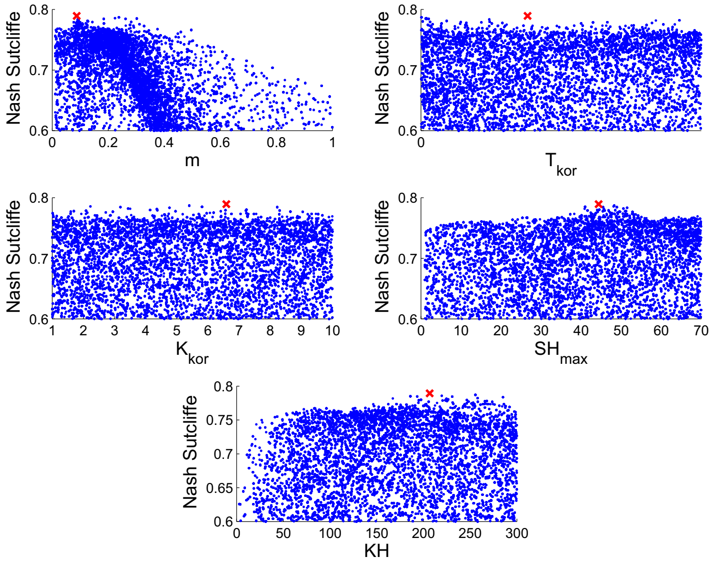


Figure B.15: The dotty plots for the Weida catchment using a 100 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

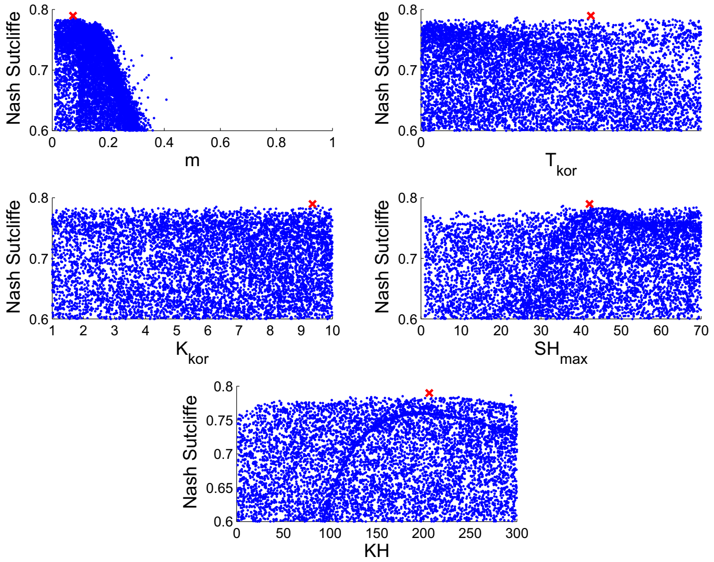


Figure B.16: The dotty plots for the Weida catchment using a 300 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

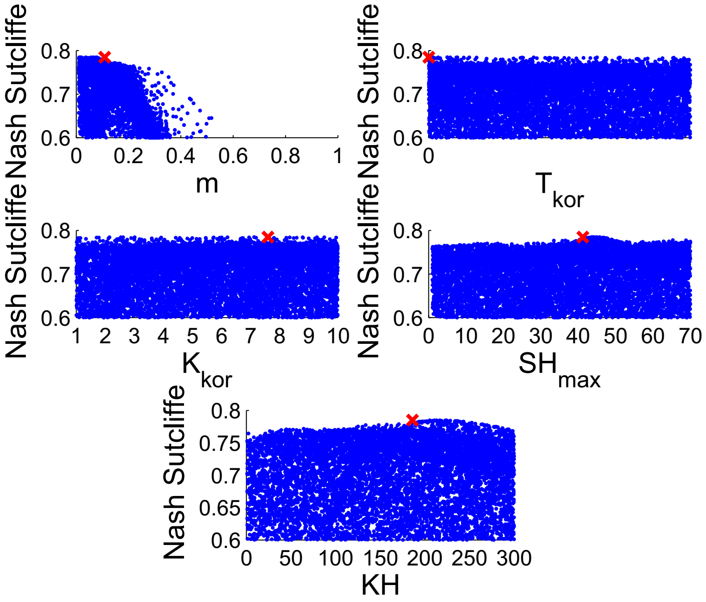


Figure B.17: The dotty plots for the Weida catchment using a 500 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

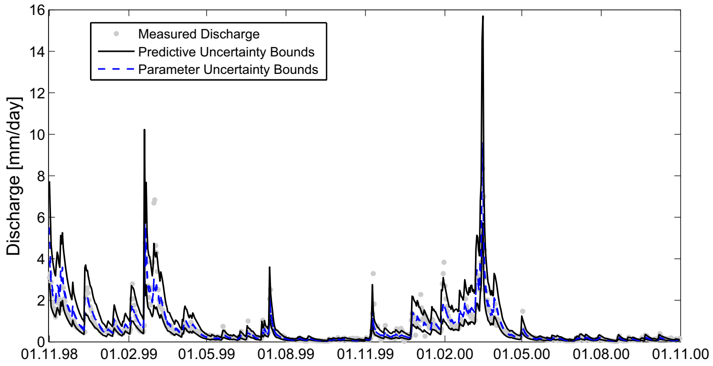


Figure B.18: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 without autocorrelation and a 100 m grid resolution.

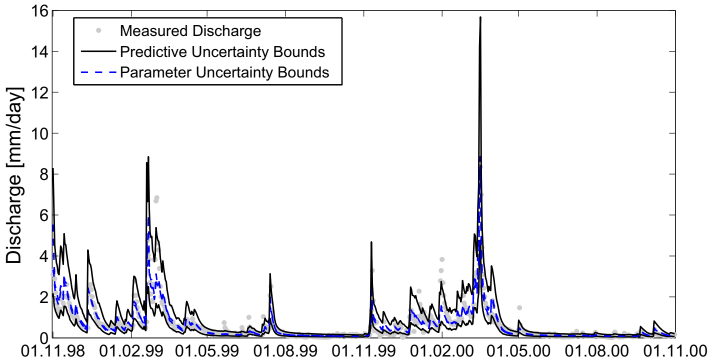


Figure B.19: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 without autocorrelation and a 300 m grid resolution.

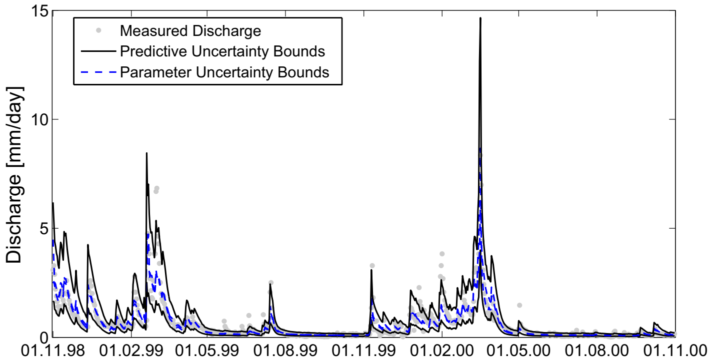


Figure B.20: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 without autocorrelation and a 500 m grid resolution.

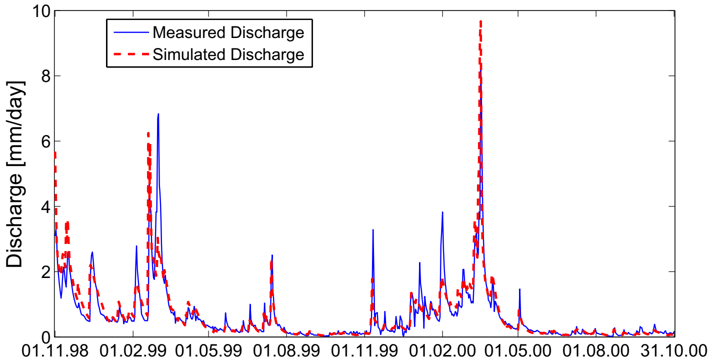


Figure B.21: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 without autocorrelation and a 100 m grid resolution.

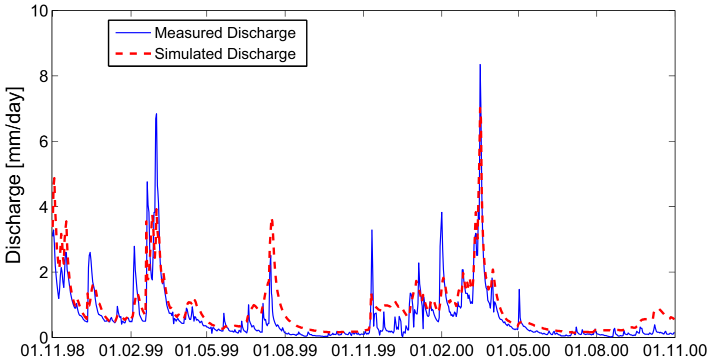


Figure B.22: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 without autocorrelation and a 300 m grid resolution.

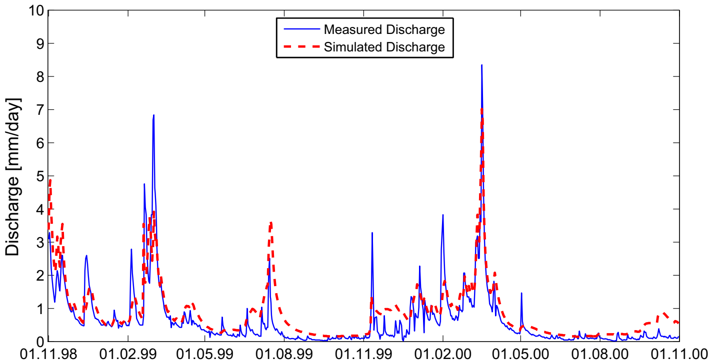


Figure B.23: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 without autocorrelation and a 500 m grid resolution.

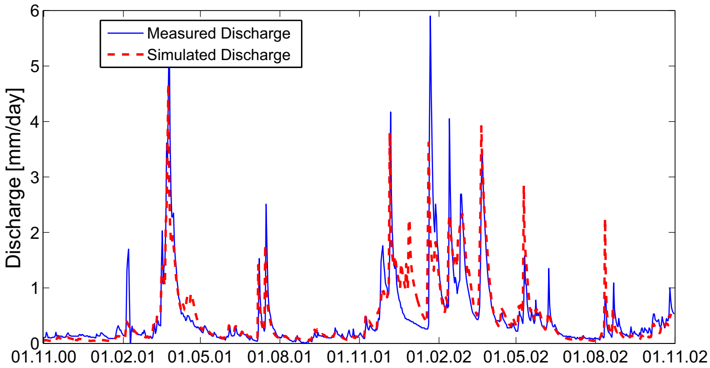


Figure B.24: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 without autocorrelation and a 100 m grid resolution for the validation period.

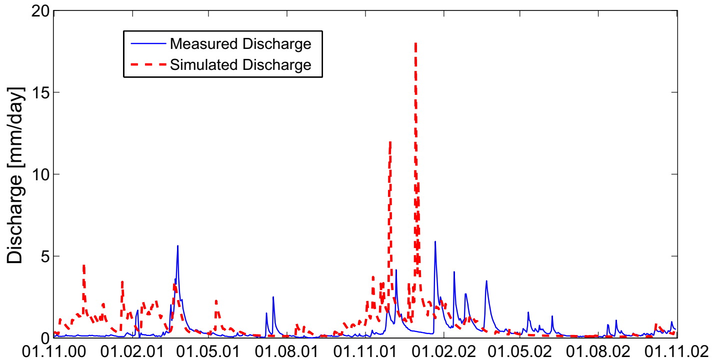


Figure B.25: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 without autocorrelation and a 200 m grid resolution for the validation period.

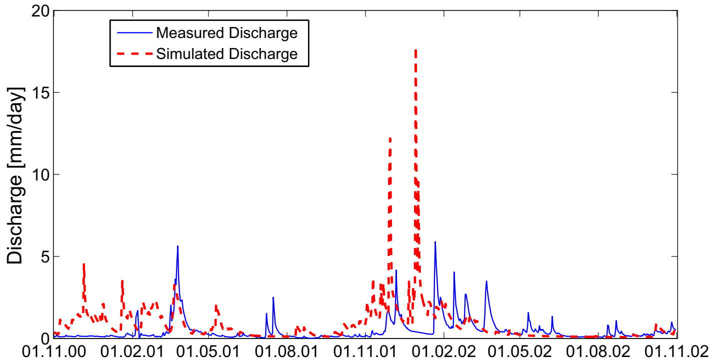


Figure B.26: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 without autocorrelation and a 300 m grid resolution for the validation period.

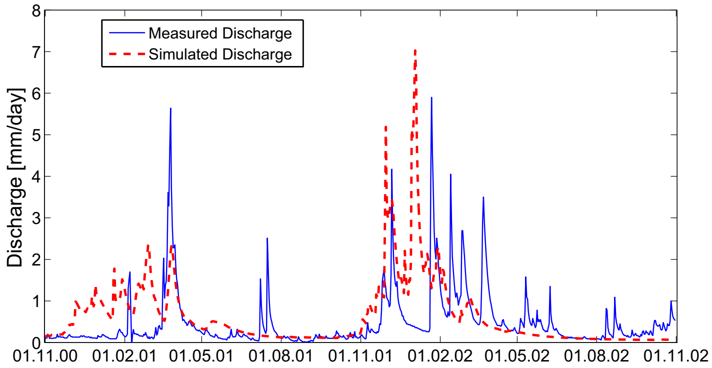


Figure B.27: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 without autocorrelation and a 500 m grid resolution for the validation period.

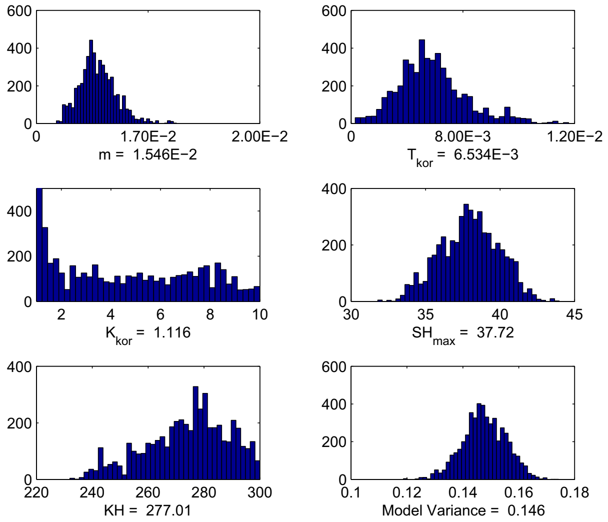


Figure B.28: The parameter posterior distribution for the Weida catchment using the MC^2 method without autocorrelation and the 100 m grid resolution.

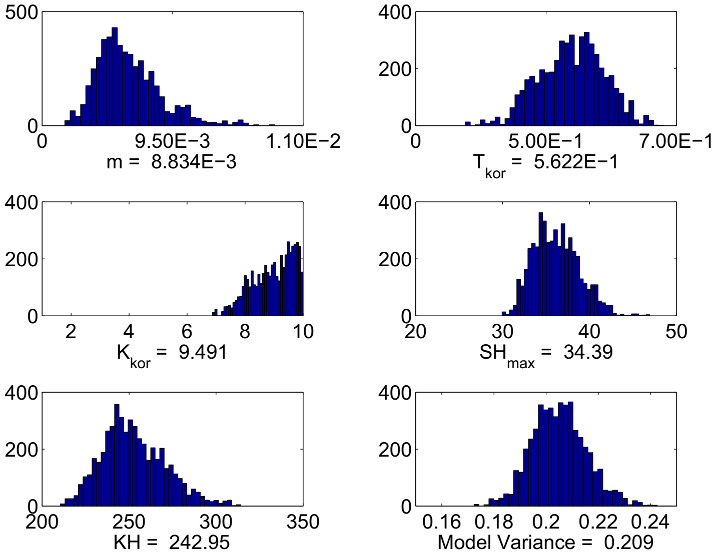


Figure B.29: The parameter posterior distribution for the Weida catchment using the MC^2 method without autocorrelation and the 200 m grid resolution.

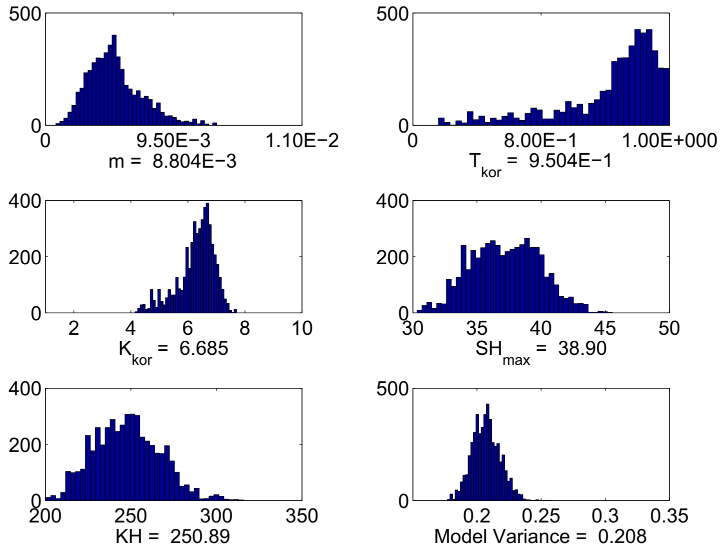


Figure B.30: The parameter posterior distribution for the Weida catchment using the MC^2 method without autocorrelation and the 300 m grid resolution.

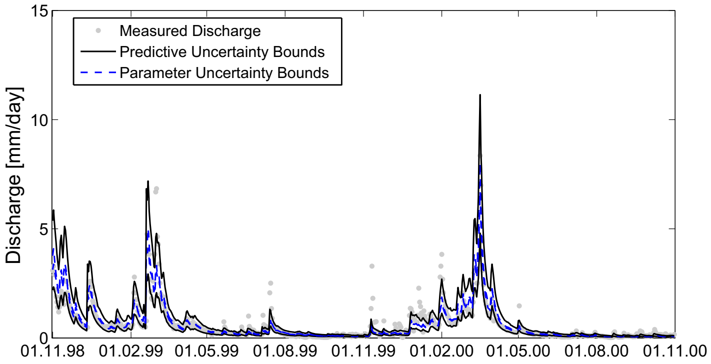


Figure B.31: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 without autocorrelation and a 200 m grid resolution.

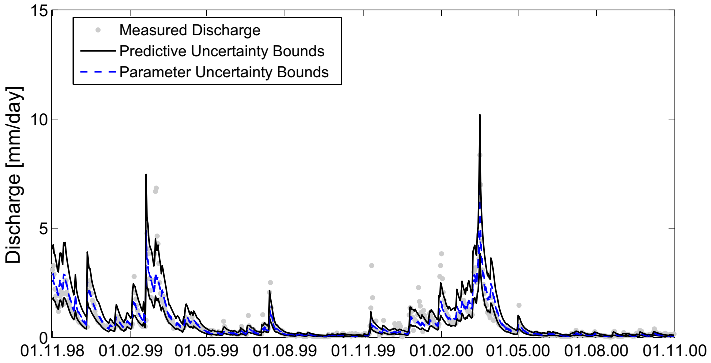


Figure B.32: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 without autocorrelation and a 300 m grid resolution.

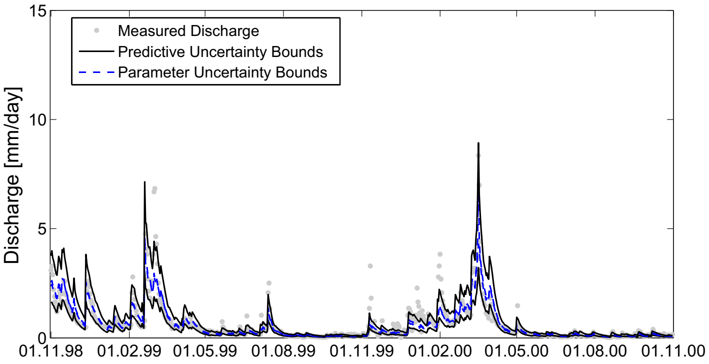


Figure B.33: The predictive and parameter uncertainty bounds for the Weida catchment using MC^2 without autocorrelation and a 500 m grid resolution.

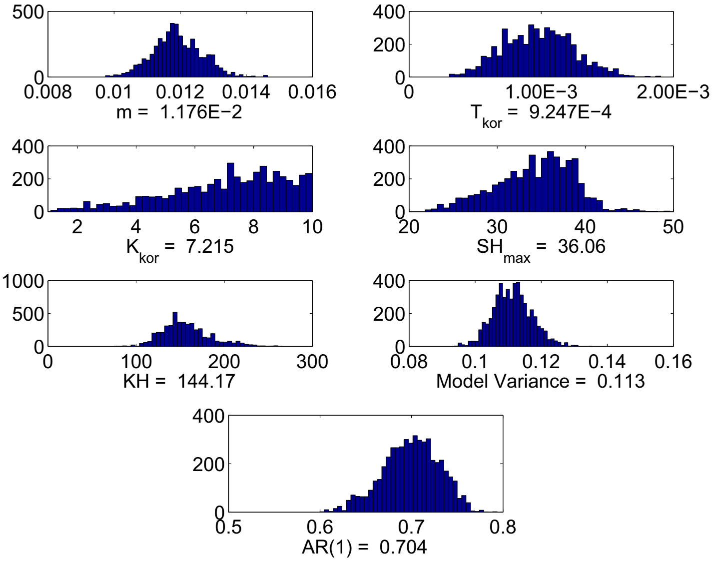


Figure B.34: The parameter posterior distribution for the Weida catchment using the MC^2 method with autocorrelation and the 200 m grid resolution.

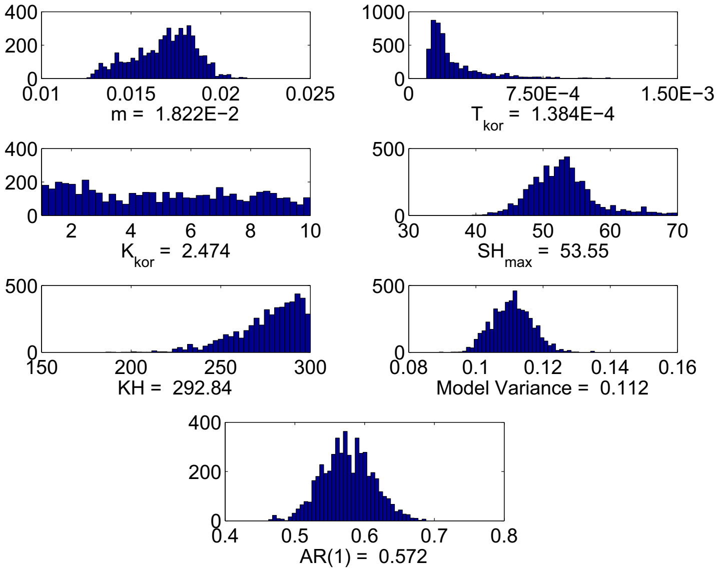


Figure B.35: The parameter posterior distribution for the Weida catchment using the MC^2 method with autocorrelation and the 300 m grid resolution.

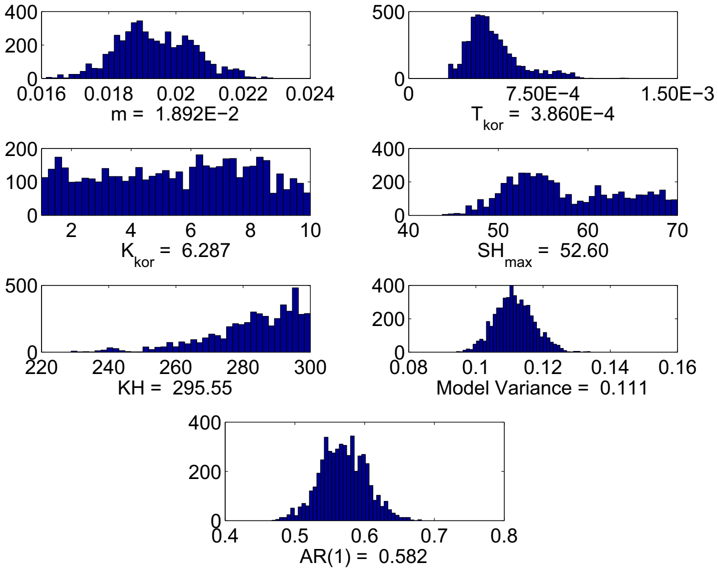


Figure B.36: The parameter posterior distribution for the Weida catchment using the MC^2 method with autocorrelation and the 500 m grid resolution.

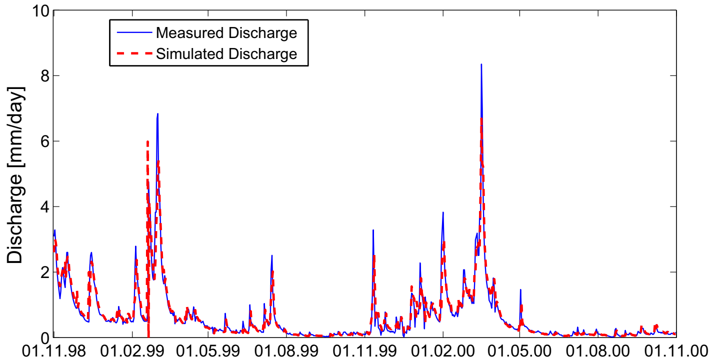


Figure B.37: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 with autocorrelation and a 200 m grid resolution.

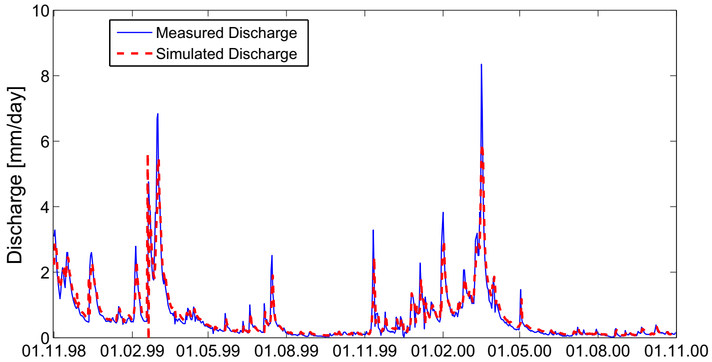


Figure B.38: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 with autocorrelation and a 300 m grid resolution.

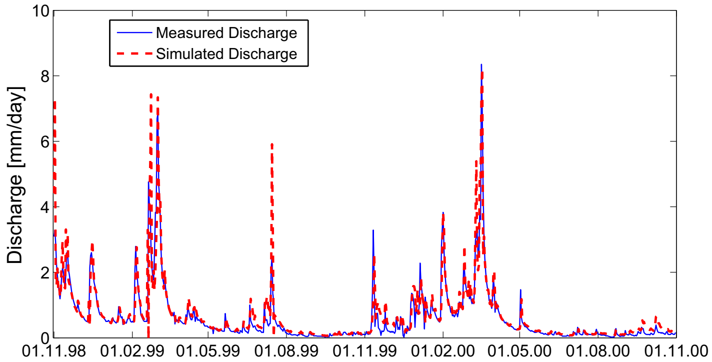


Figure B.39: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 with autocorrelation and a 500 m grid resolution.

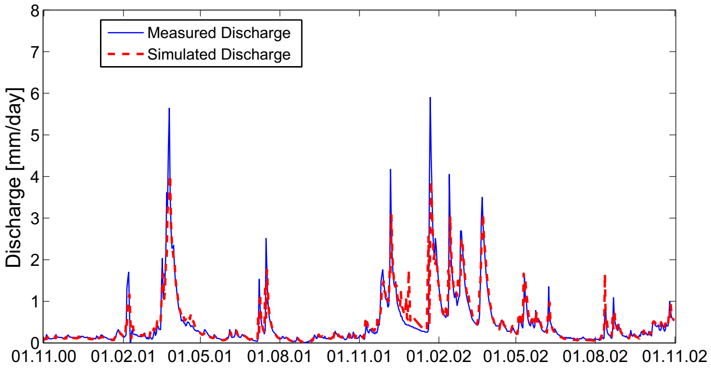


Figure B.40: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 with autocorrelation and a 100 m grid resolution.

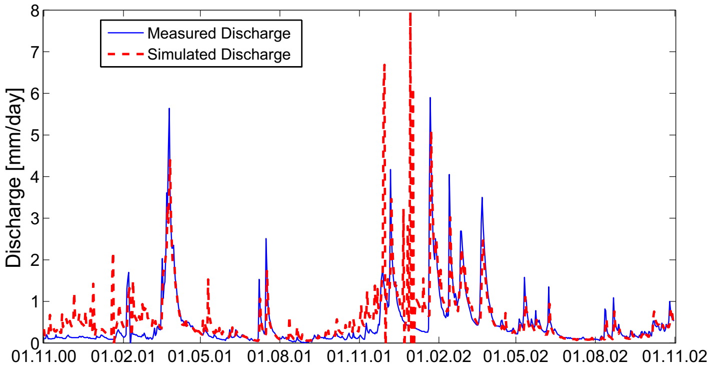


Figure B.41: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 with autocorrelation and a 200 m grid resolution.

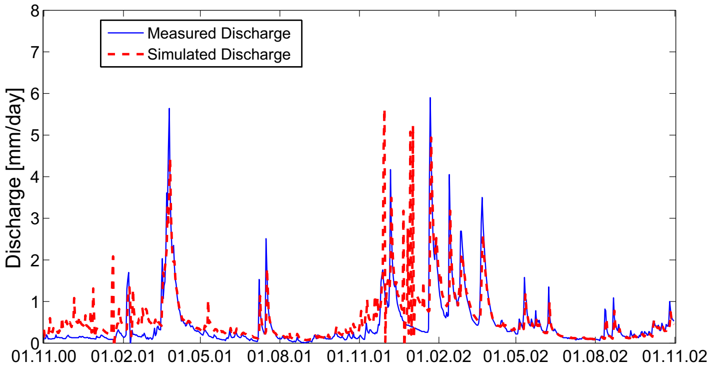


Figure B.42: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 with autocorrelation and a 300 m grid resolution for the validation period.

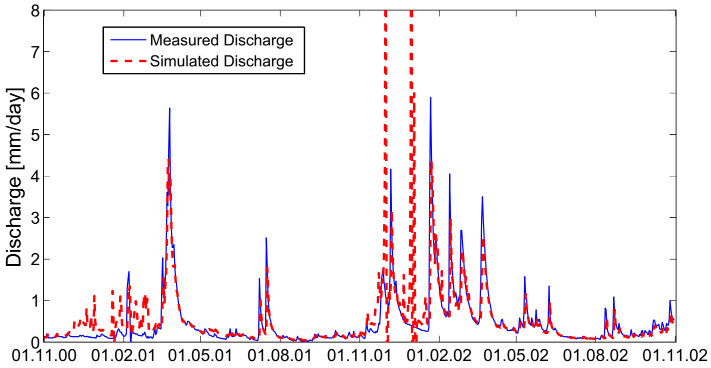


Figure B.43: The measured and simulated discharge for the Weida catchment using the best paramter set from MC^2 with autocorrelation and a 500 m grid resolution for the validation period.

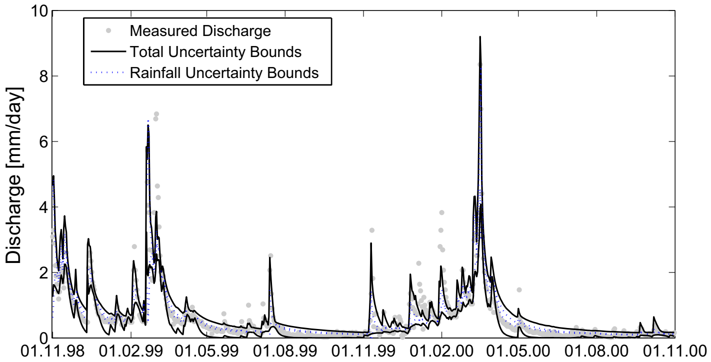


Figure B.44: The total and the rainfall uncertainty bounds for Weida catchment using INPUT concept and a 500 m grid resolution.

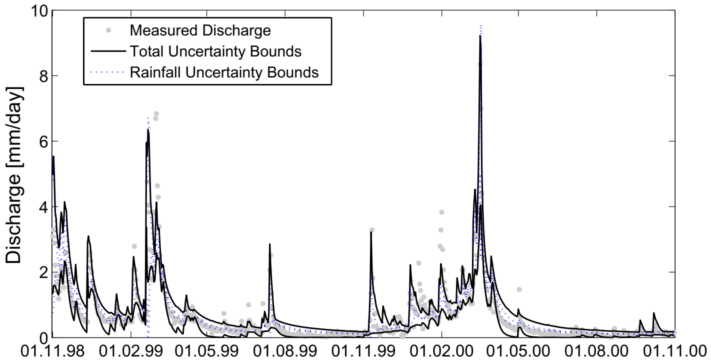


Figure B.45: The total and the rainfall uncertainty bounds for Weida catchment using INPUT concept and a 300 m grid resolution.

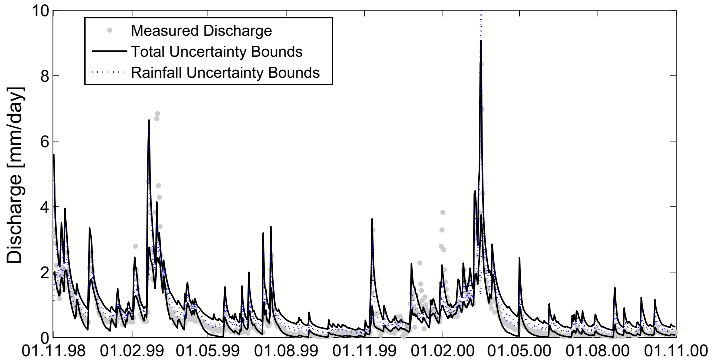


Figure B.46: The total and the rainfall uncertainty bounds for Weida catchment using INPUT concept and a 100 m grid resolution.

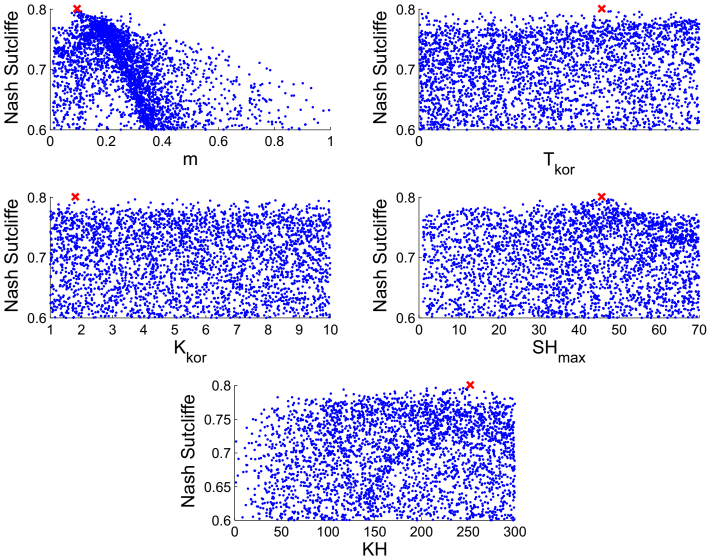


Figure B.47: The dotty plots for the Weida catchment using the INPUT concept and a 100 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

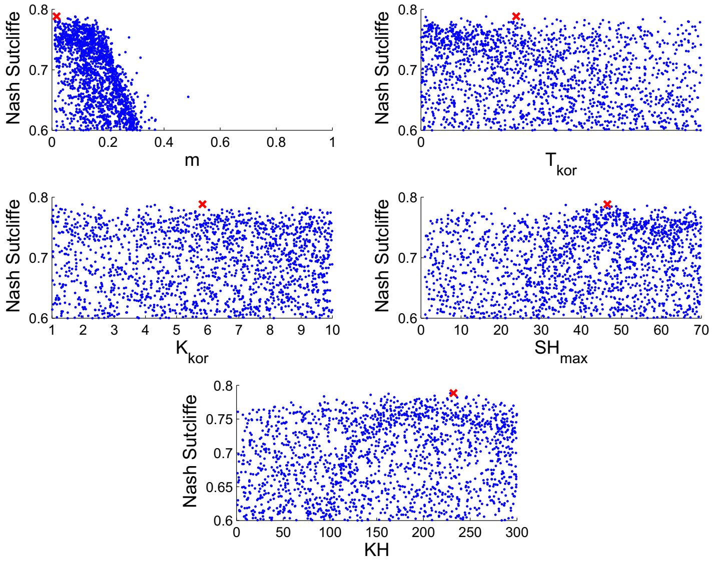


Figure B.48: The dotty plots for the Weida catchment using the INPUT concept and a 300 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

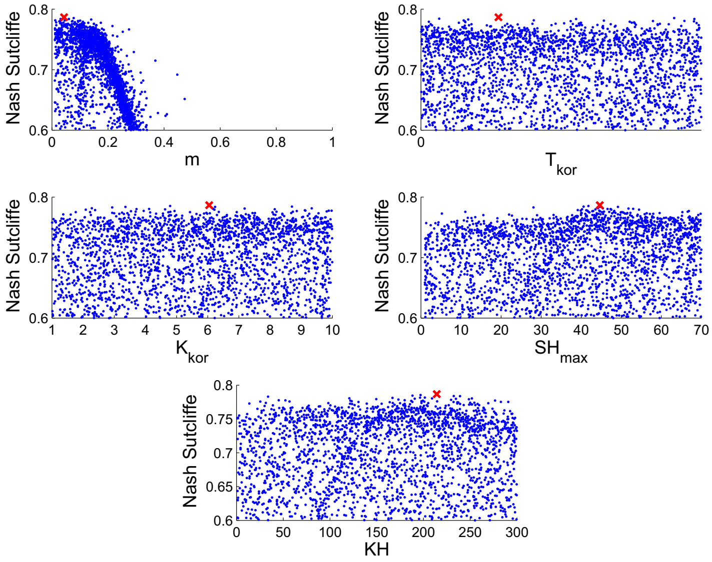


Figure B.49: The dotty plots for the Weida catchment using the INPUT concept and a 500 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

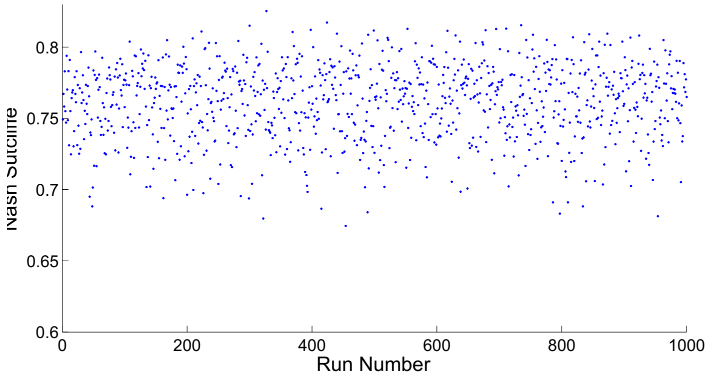


Figure B.50: The dotty plots for the Weida catchment using the INPUT concept for the second run and a 500 m grid resolution.

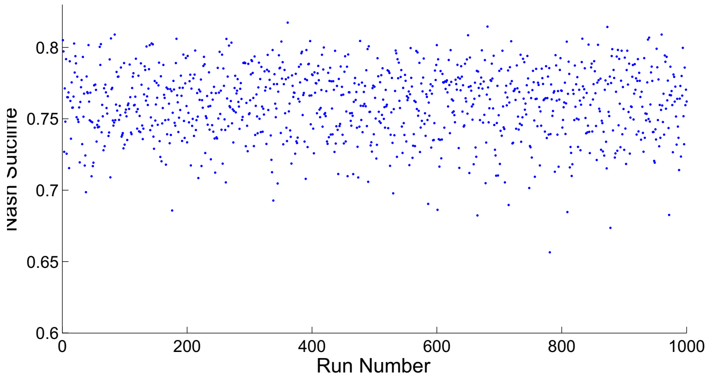


Figure B.51: The dotty plots for the Weida catchment using the INPUT concept for the second run and a 300 m grid resolution.

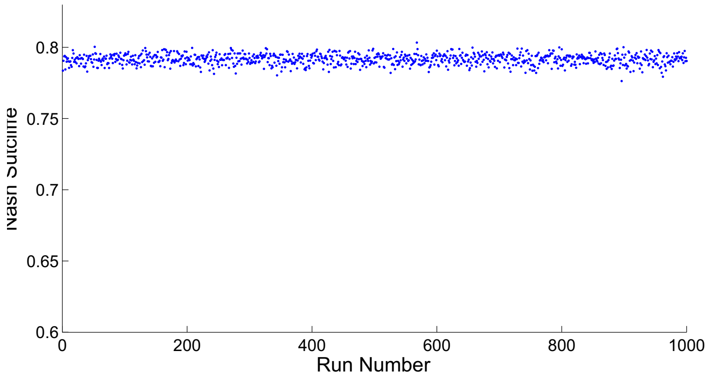


Figure B.52: The dotty plots for the Weida catchment using the INPUT concept for the second run and a 100 m grid resolution.

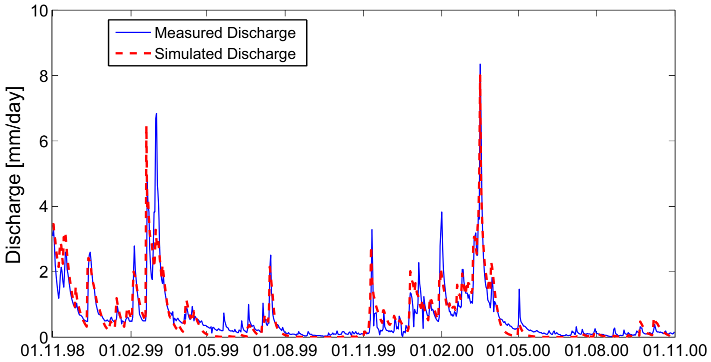


Figure B.53: The measured and simulated discharge for the Weida catchment using the best paramter set from the INPUT concept and a 500 m grid resolution.

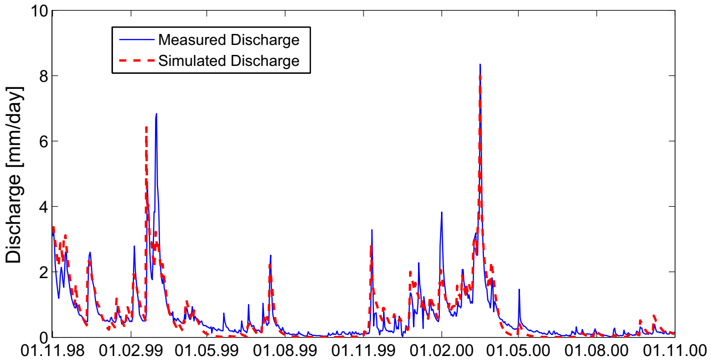


Figure B.54: The measured and simulated discharge for the Weida catchment using the best paramter set from the INPUT concept and a 300 m grid resolution.

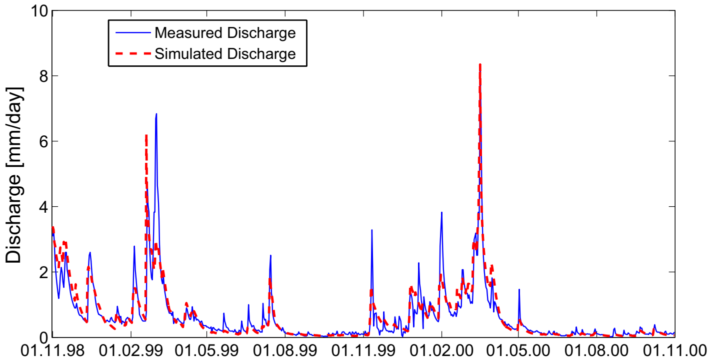


Figure B.55: The measured and simulated discharge for the Weida catchment using the best paramter set from the INPUT concept and a 100 m grid resolution.

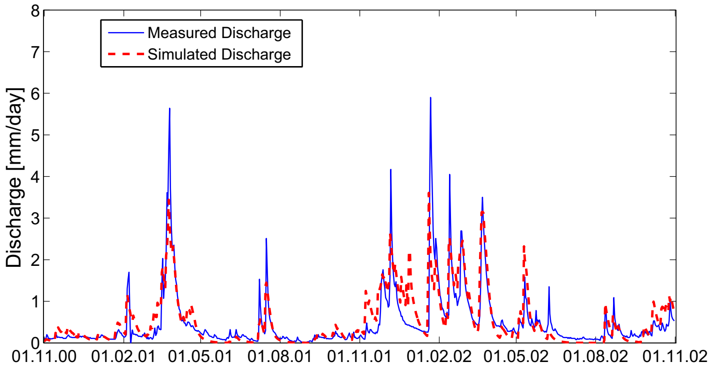


Figure B.56: The measured and simulated discharge for the Weida catchment using the best paramter set from the INPUT concept and a 500 m grid resolution for the validation period.

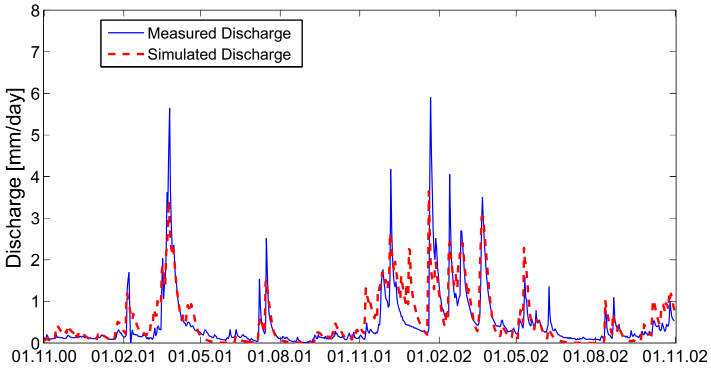


Figure B.57: The measured and simulated discharge for the Weida catchment using the best paramter set from the INPUT concept and a 500 m grid resolution for the validation period.

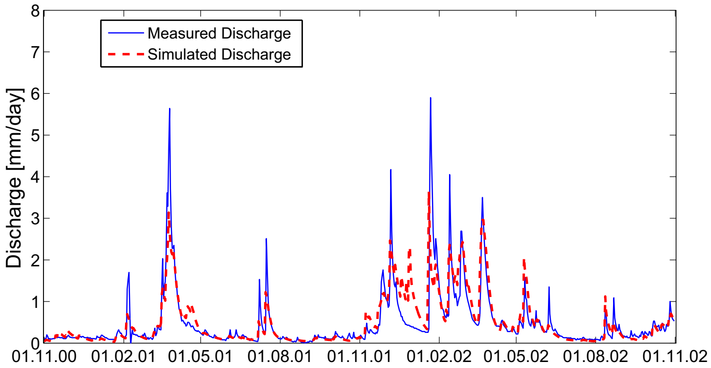


Figure B.58: The measured and simulated discharge for the Weida catchment using the best paramter set from the INPUT concept and a 500 m grid resolution for the validation period.

Appendix C

Döllnitz Figures

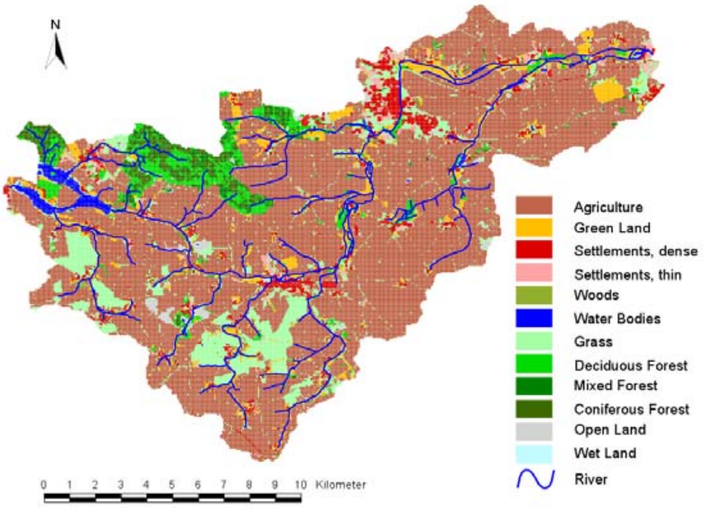


Figure C.1: The land use in the Döllnitz catchment.

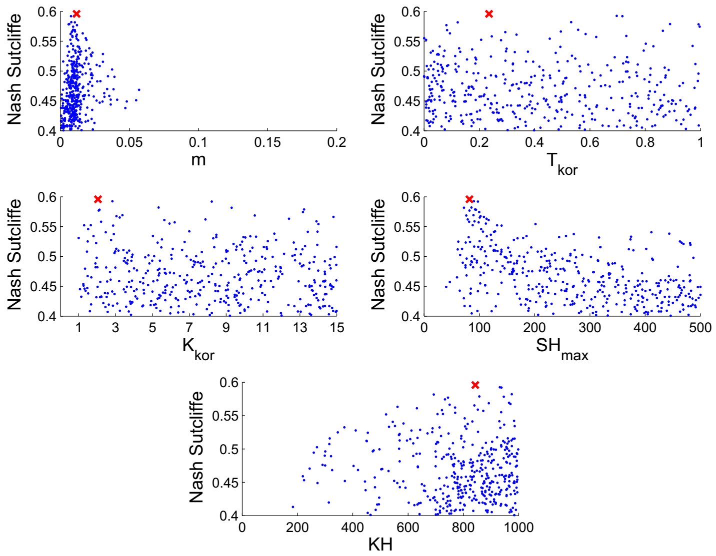


Figure C.2: The dotty plots for the Döllnitz catchment using a 100 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

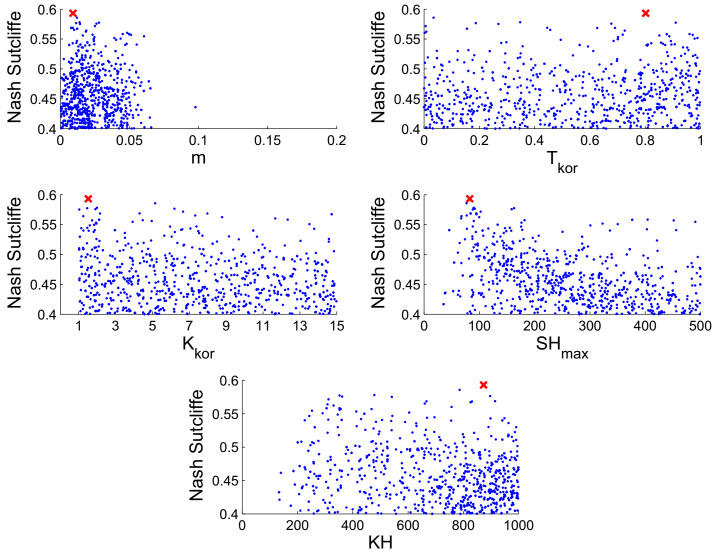


Figure C.3: The dotty plots for the Döllnitz catchment using a 500 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

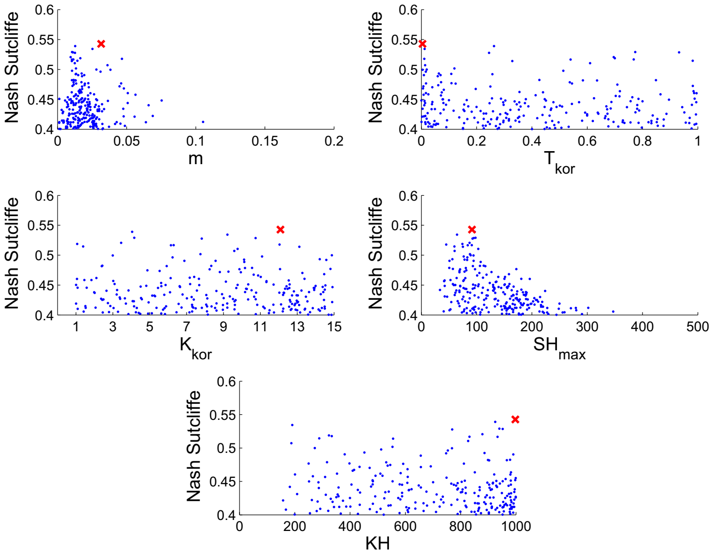


Figure C.4: The dotty plots for the Döllnitz catchment using a 1000 m grid resolution. The red cross represent the parameter with the highest Nash Sutcliffe.

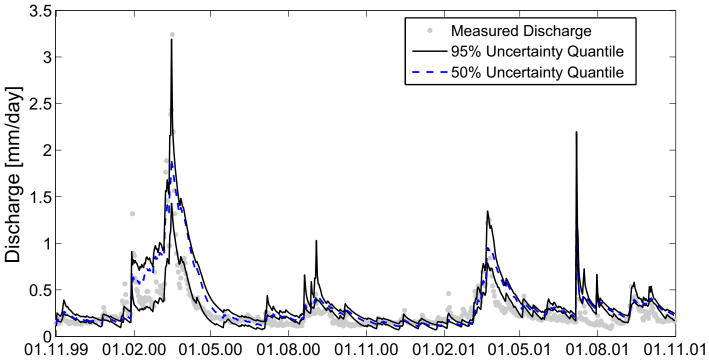


Figure C.5: The 95% and 50% uncertainty quantile for the Döllnitz catchment using the GLUE method and a 100 m grid resolution.

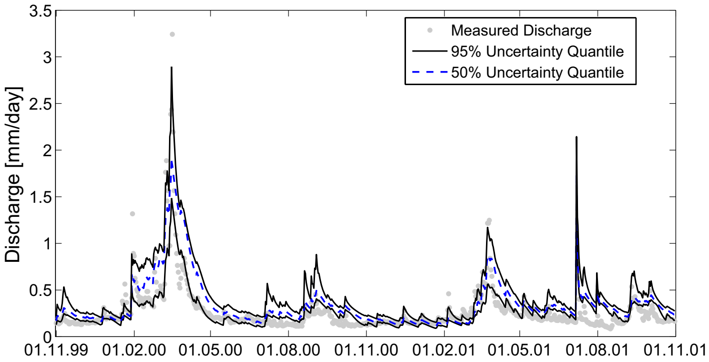


Figure C.6: The 95% and 50% uncertainty quantile for the Döllnitz catchment using the GLUE method and a 500 m grid resolution.

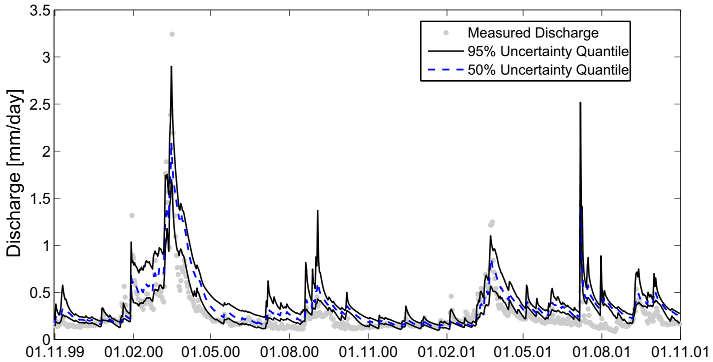


Figure C.7: The 95% and 50% uncertainty quantile for the Döllnitz catchment using the GLUE method and a 1000 m grid resolution.

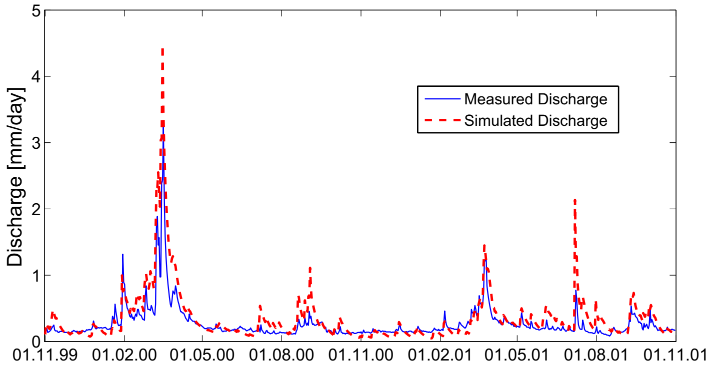


Figure C.8: The measured and simulated discharge for the Döllnitz catchment using the best paramter set from GLUE and a 300 m grid resolution.

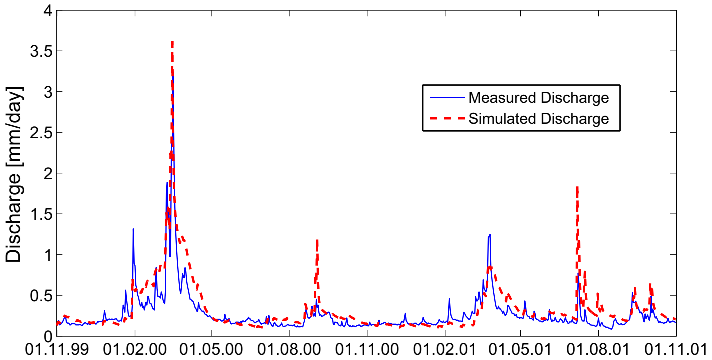


Figure C.9: The measured and simulated discharge for the Döllnitz catchment using the best paramter set from GLUE and a 500 m grid resolution.

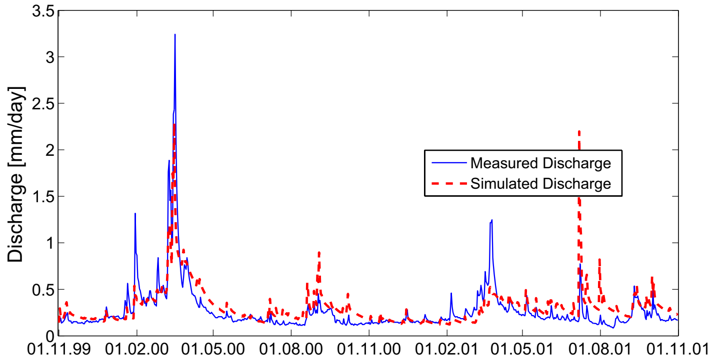


Figure C.10: The measured and simulated discharge for the Döllnitz catchment using the best paramter set from GLUE and a 1000 m grid resolution.

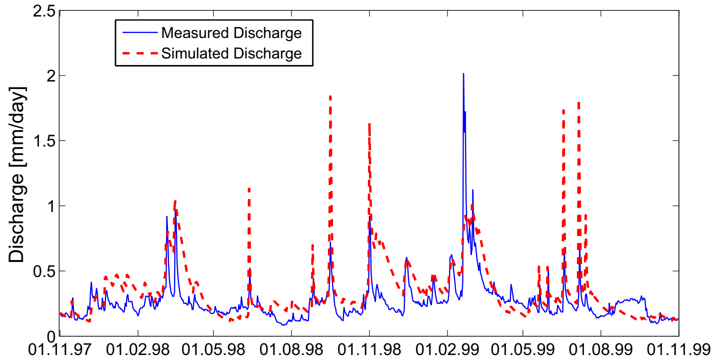


Figure C.11: The measured and simulated discharge for the Döllnitz catchment using the best paramter set from GLUE and a 100 m grid resolution for the validation period.

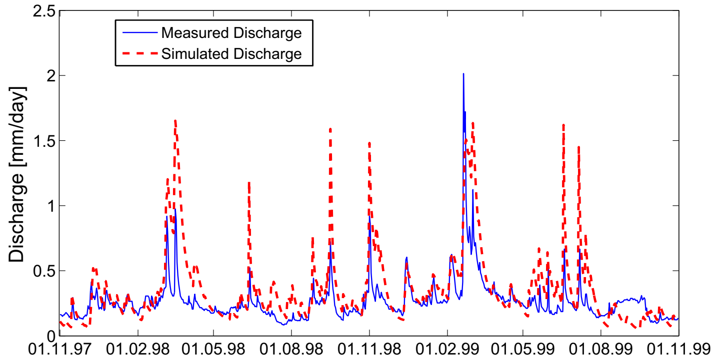


Figure C.12: The measured and simulated discharge for the Döllnitz catchment using the best paramter set from GLUE and a 300 m grid resolution for the validation period.

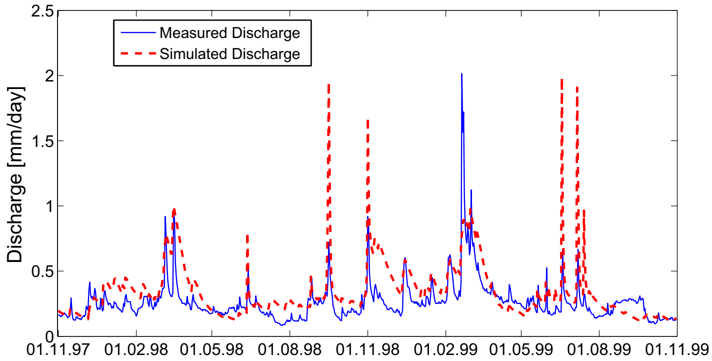


Figure C.13: The measured and simulated discharge for the Döllnitz catchment using the best paramter set from GLUE and a 500 m grid resolution for the validation period.

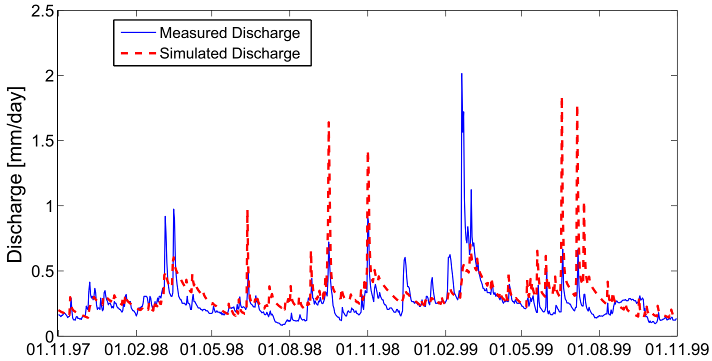


Figure C.14: The measured and simulated discharge for the Döllnitz catchment using the best paramter set from GLUE and a 1000 m grid resolution for the validation period.

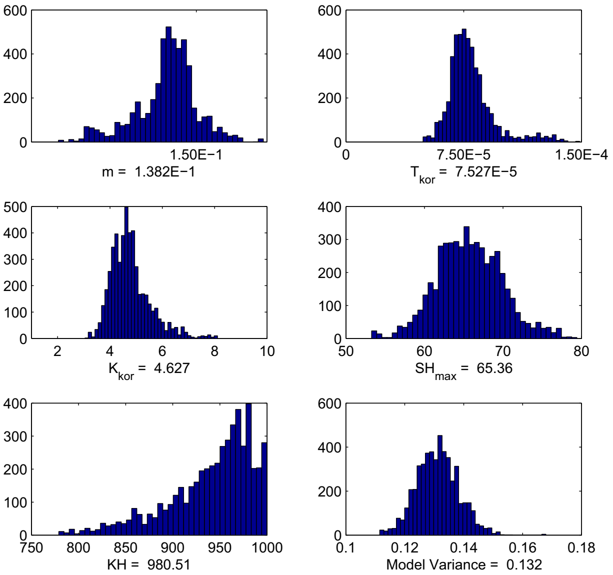


Figure C.15: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method without autocorrelation and the 300 m grid resolution.

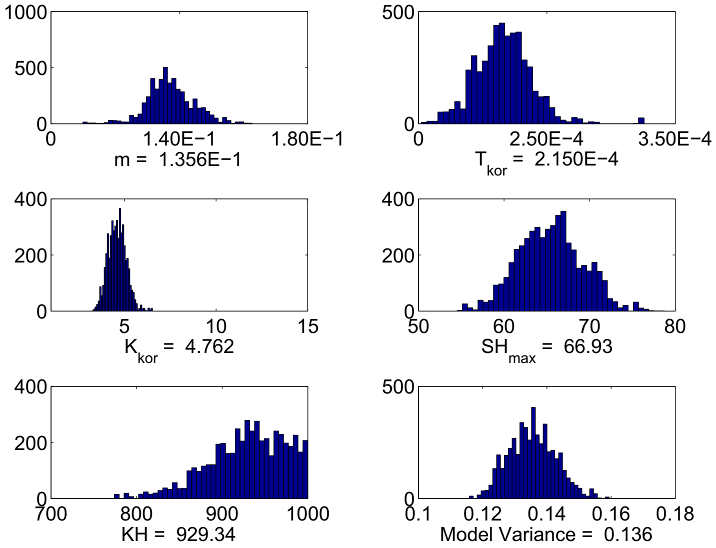


Figure C.16: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method without autocorrelation and the 500 m grid resolution.

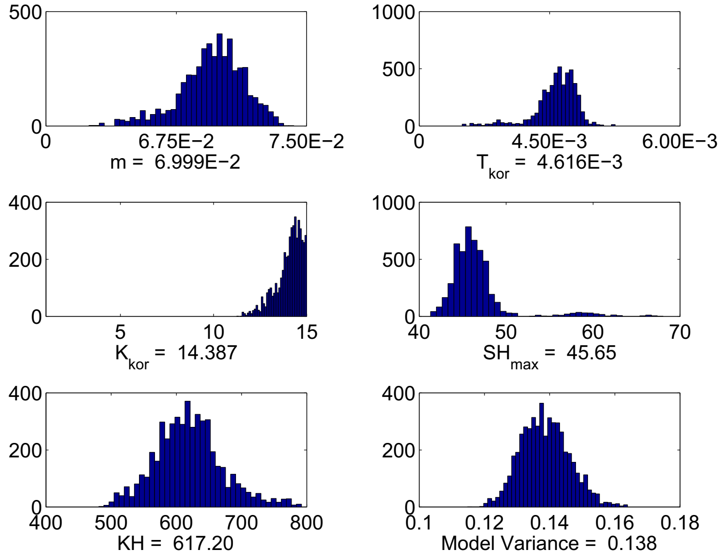


Figure C.17: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method without autocorrelation and the 1000 m grid resolution.

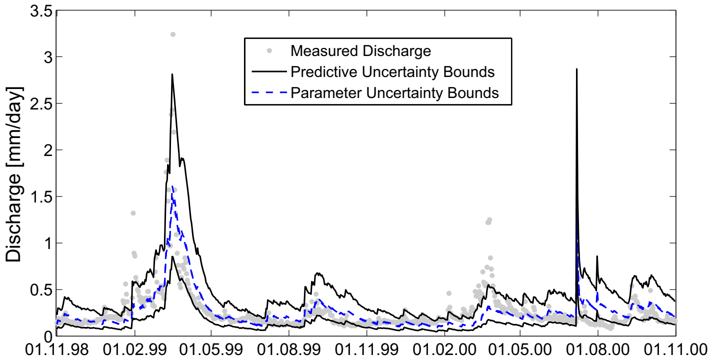


Figure C.18: The predictive and parameter uncertainty bounds for the Döllnitz catchment using MC^2 without autocorrelation and a 100 m grid resolution.

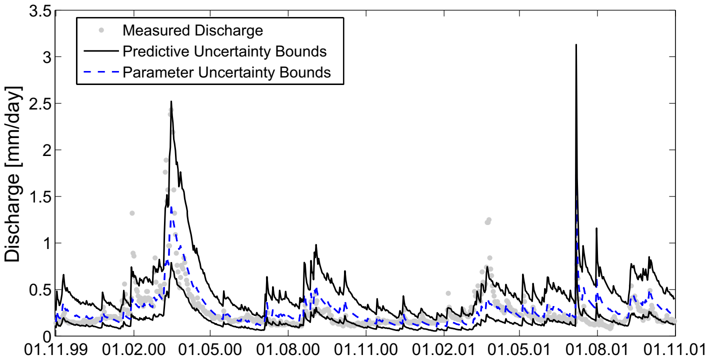


Figure C.19: The predictive and parameter uncertainty bounds for the Döllnitz catchment using MC^2 without autocorrelation and a 300 m grid resolution.

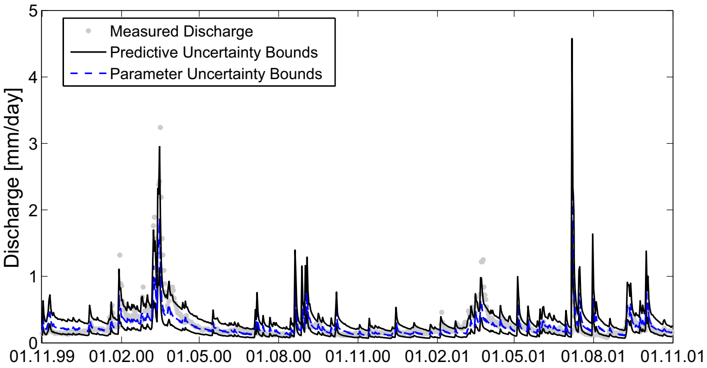


Figure C.20: The predictive and parameter uncertainty bounds for the Döllnitz catchment using MC^2 without autocorrelation and a 1000 m grid resolution.

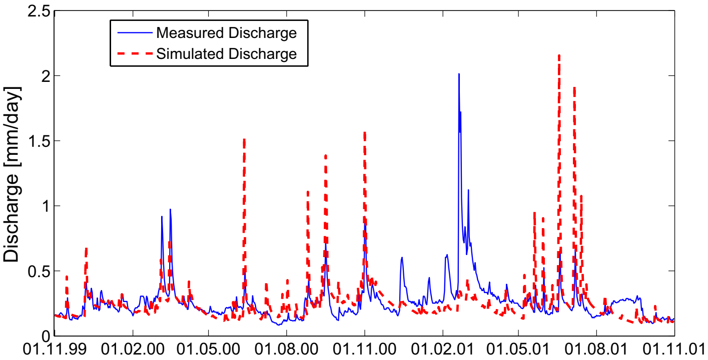


Figure C.21: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 100 m grid.

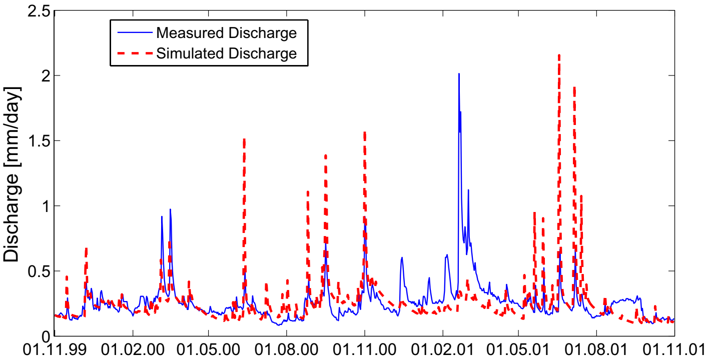


Figure C.22: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 300 m grid.

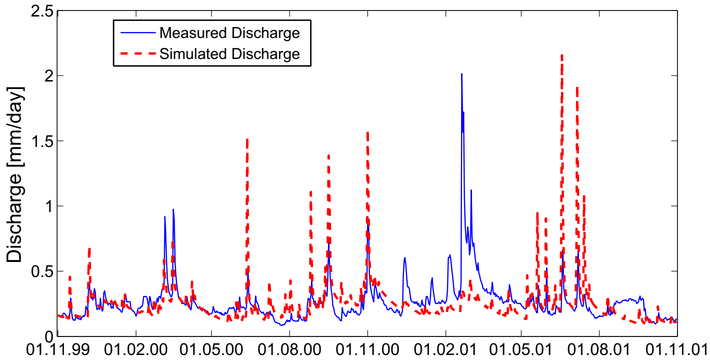


Figure C.23: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 500 m grid resolution in the validation period.

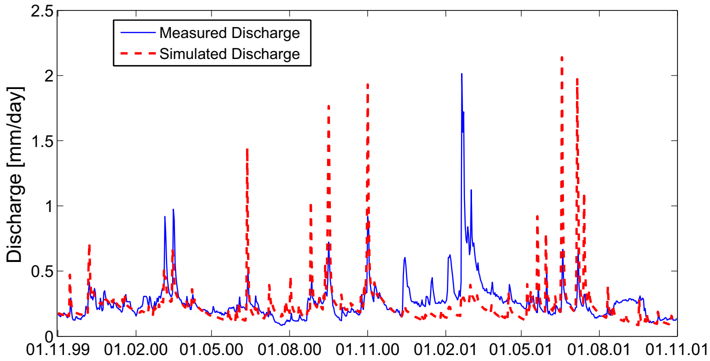


Figure C.24: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 1000 m grid resolution in the validation period.

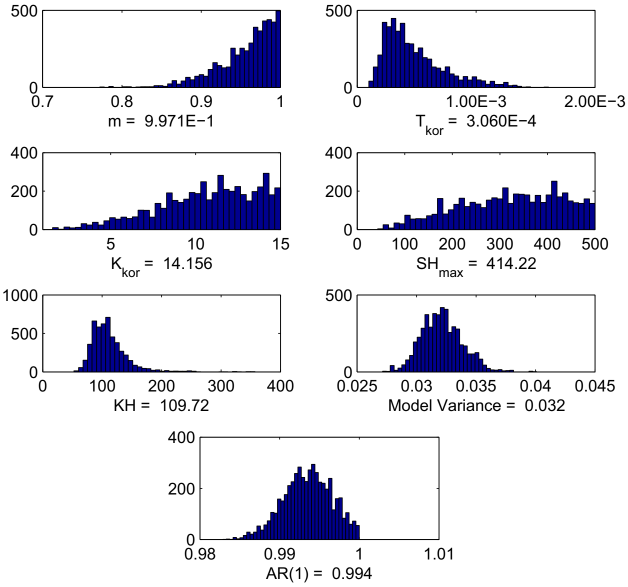


Figure C.25: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method with autocorrelation and the 100 m grid resolution.

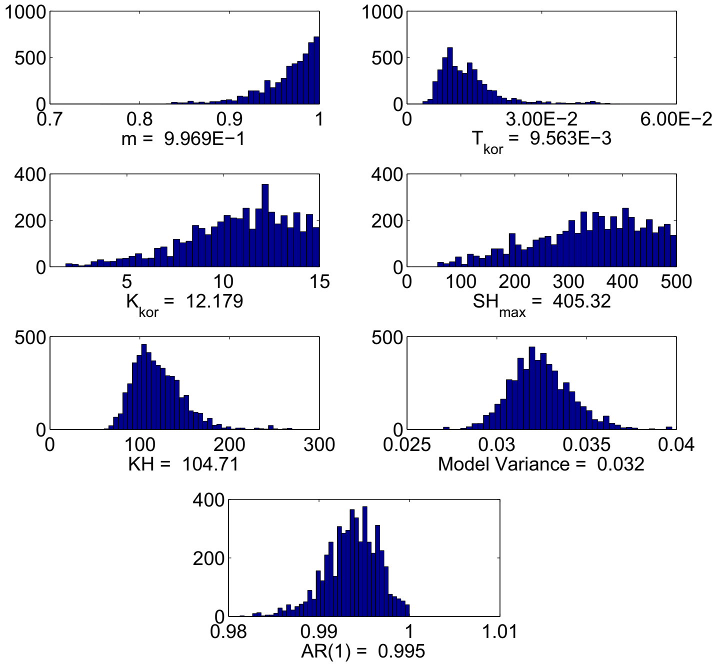


Figure C.26: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method with autocorrelation and the 500 m grid resolution.

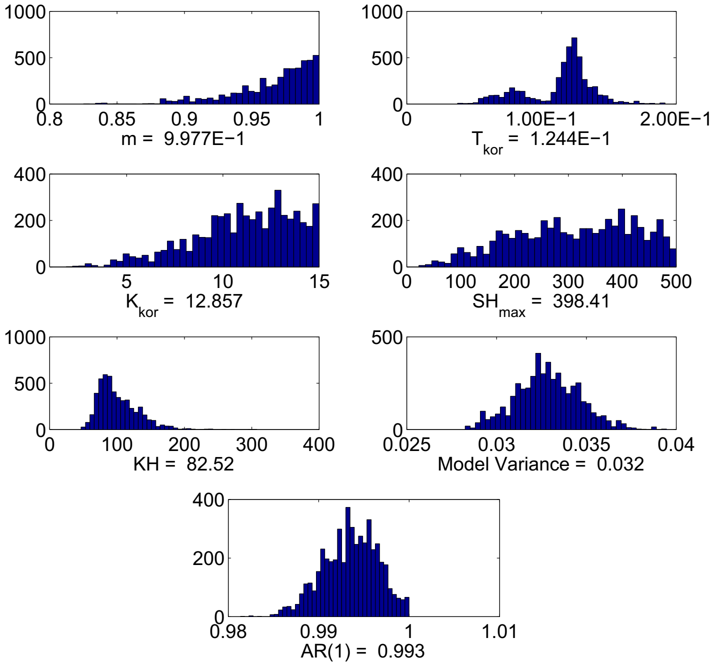


Figure C.27: The parameter posterior distribution for the Döllnitz catchment using the MC^2 method with autocorrelation and the 1000 m grid resolution.

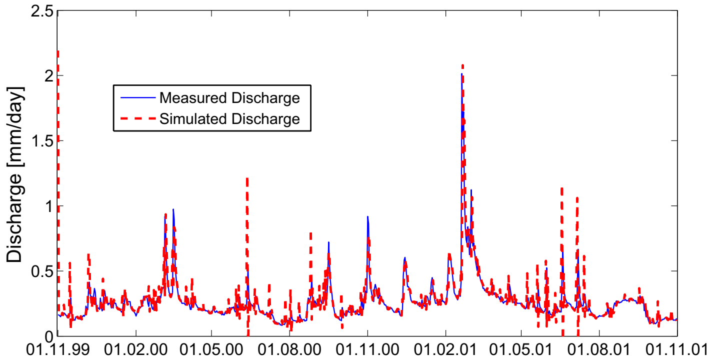


Figure C.28: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 100 m grid resolution in the validation period.

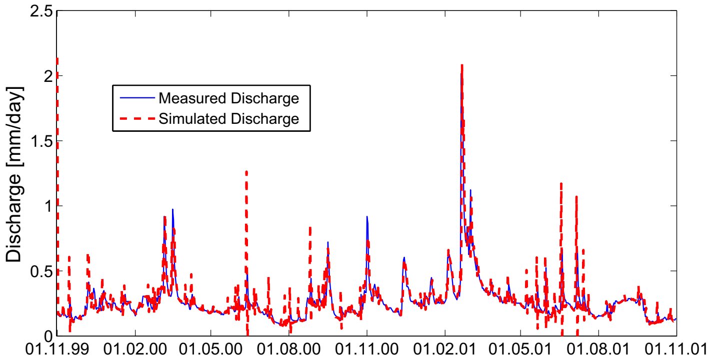


Figure C.29: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 300 m grid resolution in the validation period.

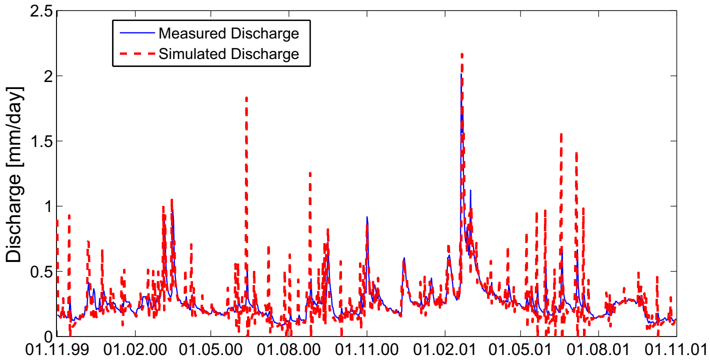


Figure C.30: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 500 m grid resolution in the validation period.

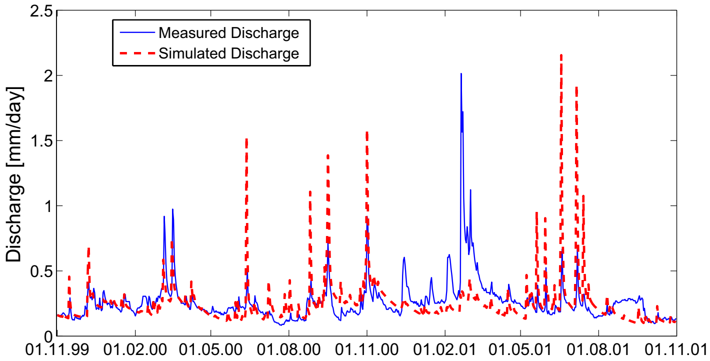


Figure C.31: The measured and simulated discharge for the Döllnitz catchment using the best parameter set from MC^2 with autocorrelation and a 1000 m grid resolution in the validation period.

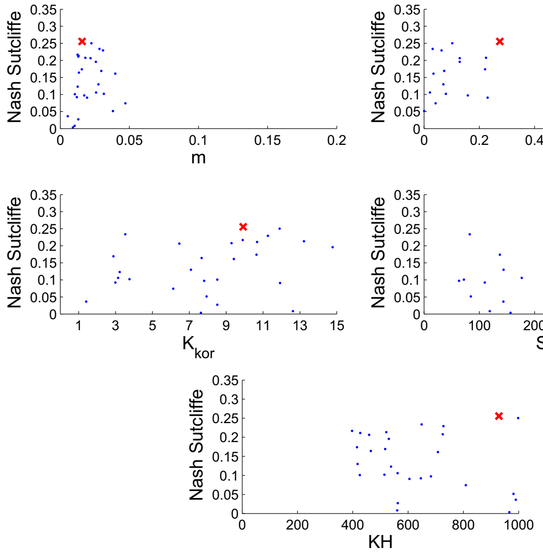


Figure C.32: Dotty Plots for the 100 m Grid Resolution for Doellnitz unsing the New konzept.

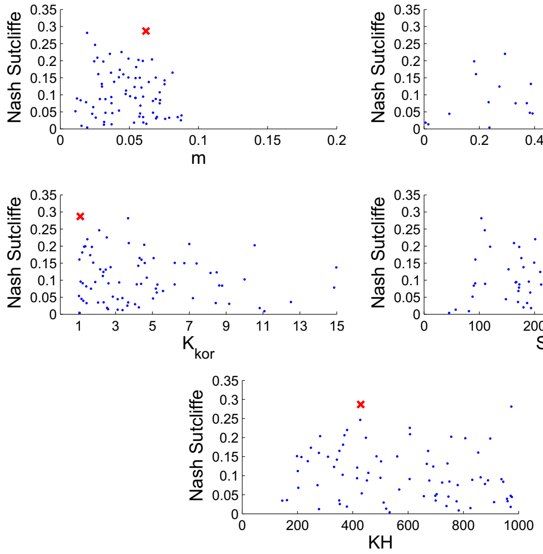


Figure C.33: Dotty Plots for the 500 m Grid Resolution for Doellnitz unsing the New konzept.

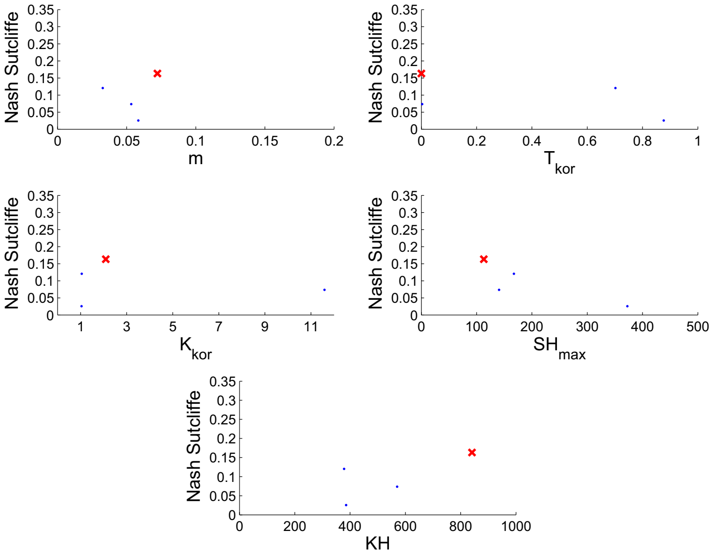


Figure C.34: Dotty Plots for the 1000 m Grid Resolution for Doellnitz unsing the New konzept.

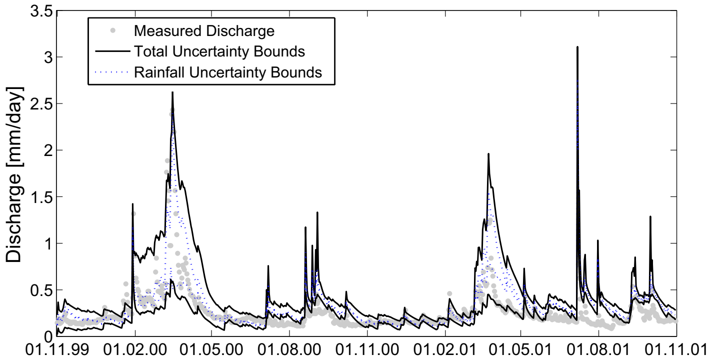


Figure C.35: The Uncertainty Quantiles based on the new konzept Methodology for the Doellnitz Catchment using a 100 m Grid Resolution

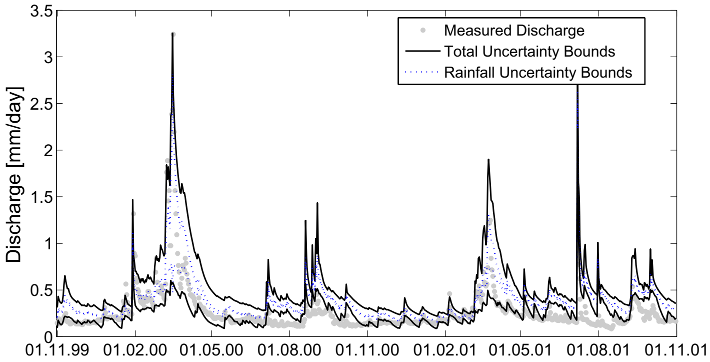


Figure C.36: The Uncertainty Quantiles based on the new konzept Methodology for the Doellnitz Catchment using a 500 m Grid Resolution

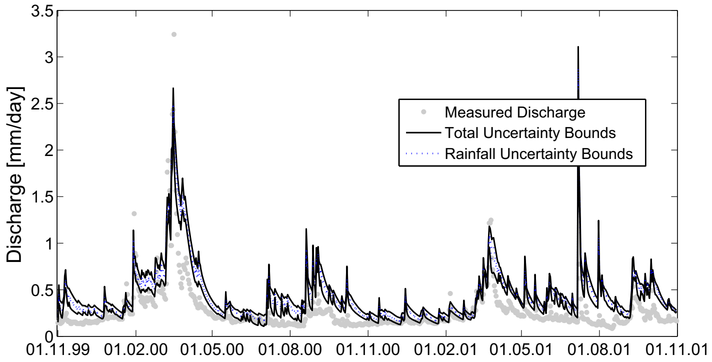


Figure C.37: The Uncertainty Quantiles based on the new konzept Methodology for the Doellnitz Catchment using a 1000 m Grid Resolution

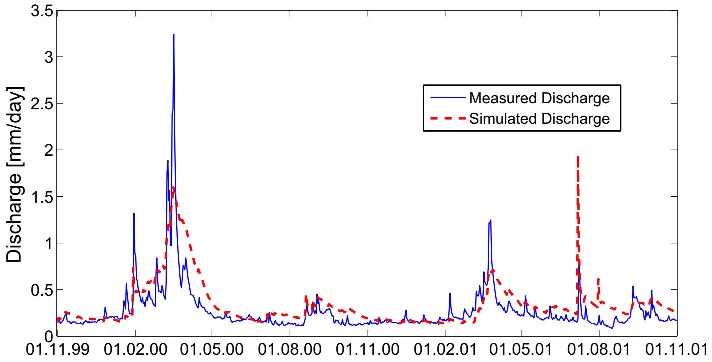


Figure C.38: The Measured and Simulated Discharge for the Doellnitz Catchment using the new konzept with a 100 m Grid Resolution.

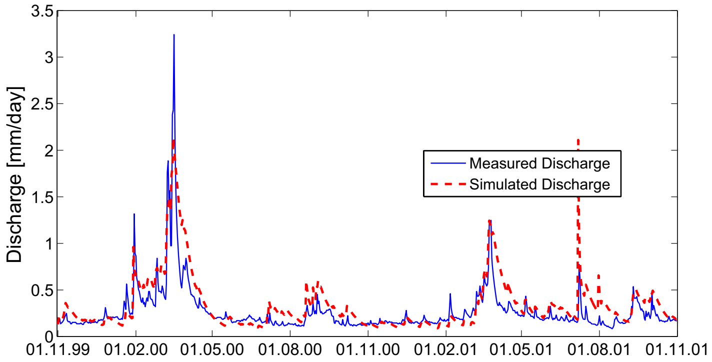


Figure C.39: The Measured and Simulated Discharge for the Doellnitz Catchment using the new konzept with a 500 m Grid Resolution.

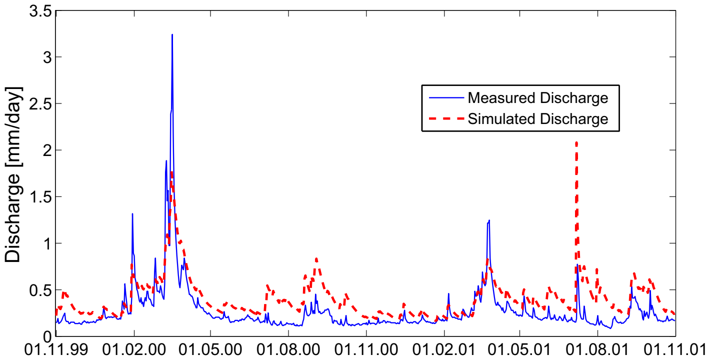


Figure C.40: The Measured and Simulated Discharge for the Doellnitz Catchment using the new konzept with a 1000 m Grid Resolution.

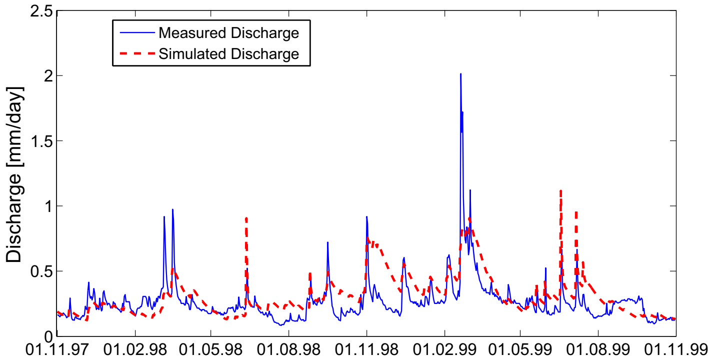


Figure C.41: The Measured and Simulated Discharge for the Doellnitz Catchment using the new konzept with a 100 m Grid Resolution for the Validation Period.

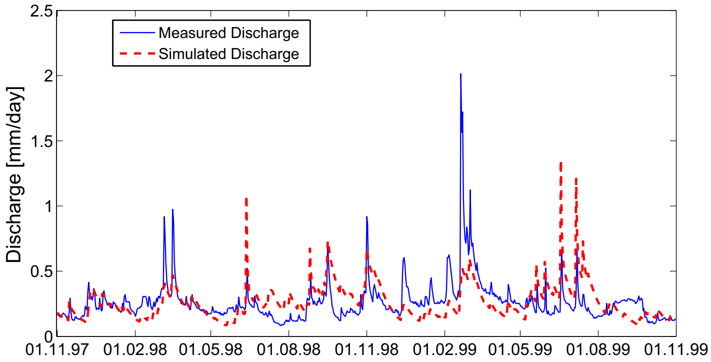


Figure C.42: The Measured and Simulated Discharge for the Doellnitz Catchment using the new konzept with a 500 m Grid Resolution for the Validation Period.

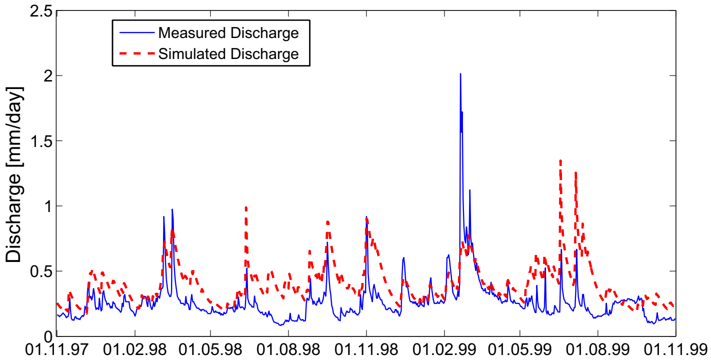


Figure C.43: The Measured and Simulated Discharge for the Doellnitz Catchment using the new konzept with a 1000 m Grid Resolution for the Validation Period.

List of Symbols and Abbreviations

Abbreviation	Unit	Description
c_p	KJ/(Kg.K)	Specific Heat Capacity of Dry Air at Constant Pressure
c_s		topographic index
CDF		Cumulative Density Function
COV		Covariance-Matrix
d(u,u$_j$)		distance to the station
E		Expected Value
E	mm/m	Latent Heat Flux
e	hPa	actual vapor Pressure
e_s	hPa	saturation vapor Pressure at Temperature T
G	Wh/m	Soil Heat Flux
GLUE		Generalized Likelihood Uncertainty Estimator
I		Input Data
IDW		Inverse Distance Weighting Interpolation
L		Likelihood Function
LfUG		saechsisches Landesamt fuer Umwelt und Geologie
m	mm	Model Parameter
M		Hydrological Model
MC		Monte Carlo (Method)
MCMC / MC^2		Monte Carlo Markov Chain
n		Number of Observation
N(μ, σ)		Normal Distribution
NS		Nash Sutcliffe Criterion
p		weighting power
P()		Probability
P(|)		Conditioned Probability
PDF		Probability Distribution Function
q	[mm/day]	Measured Discharge
r		Acceptance Rule

Abbreviation	Unit	Description
r_a	s/m	Bulk-aerodynamic Resistance
R_N		Net Radiation
r_s	s/m	Bulk-surface Resistance
s		Sensitivity Index
S_m	mm	Mean saturation deficit of the Basin
SAE		Sum of Absolute Error
SSLR		Sum of Square Logarithmic Residuals
SSR		Sum of Square Residuals
t	Second/Day	Time Step
T	C	Temperature
T_0	m/s	saturated local Transmissivity
u_{ref}		Reference Parameter Values
u^*		real Parameter Values, but improbably known
V		Variance
V_j		Variance due to x_j
w		Weighting Factors
WaSiM		Water Balance Simulation Model-ETH
z_u		interpolated value at location u
$\hat{z}(u_j)$		observed value at station j

Symbol	Unit	Description
α		Shape Factor Gamma Distribution
α		Box-Cox Transformation Factor
α_t	m/m	specific catchment area per unit length
β		Shape Factor Gamma Distribution
β_t	m/m	Slope Angel
Δ	hPa/K	Tangent of the Saturated Vapor Pressure Curve
ϵ		Error
Γ		Gamma Distribution
γ		mean topographic index of the catchment
γ_p	hPa/K	psychrometric constant
λ		Box-Cox Transformation Factor
λ	kJ/kg	Latent Vaporization Heat
μ		Mean
θ		Model Parameters
ρ	Kg/m	Density of Dry Air
ρ		First Order Autocorrelation Factor
σ		Model Variance
σ_{log}		Logarithmic Model Variance

List of Figures

List of Tables

Acknowledgements

... some text ...